Bivariate Scoring Rules: Unifying the Characterizations of Positional Scoring Rules and Kemeny's Rule

Patrick Lederer*
Technical University of Munich

April 21, 2024

This paper studies social preference functions (SPFs), which map the voters' ordinal preferences over a set of alternatives to a non-empty set of strict rankings over the alternatives. Maybe the most prominent SPFs are positional scoring rules and Kemeny's rule. While these two types of rules behave intuitively quite differently, they are axiomatically surprisingly similar and we thus provide a joint axiomatic characterization of these SPFs. To this end, we introduce the class of bivariate scoring rules which generalize Kemeny's rule by weighting comparisons between alternatives depending on their positions in the voters' preference relations. In particular, this class contains both Kemeny's rule and all positional scoring rules. As our main result, we then characterize the set of bivariate scoring rules as the only SPFs that satisfy—aside from some standard axioms—mild consistency conditions for variable electorates and variable agendas. As corollaries of this result, we also derive variants of the well-known characterizations of positional scoring rules by Smith (1973) and of Kemeny's rule by Young (1988). Thus, our result unifies the independent streams of research on positional scoring rules and Kemeny's rule by giving a joint axiomatic basis of these SPFs.

Keywords: Social Preference Functions, Positional Scoring Rules, Kemeny's Rule, Reinforcement, Agenda Consistency

JEL Classification Code: D7

*Email address: `ledererp@in.tum.de`

1. Introduction

One of the most classic problems in social choice theory is ranking aggregation: given the voters' strict preferences over a set of alternatives, a collective preference relation should be found. Indeed, even Arrow's impossibility theorem, which is commonly considered as the starting point of modern social choice theory, has originally been stated in terms of ranking aggregation (Arrow, 1951). Moreover, the problem of ranking aggregation has found its way to numerous fields, such as metasearch (Dwork et al., 2001; Renda and Straccia, 2003), computational biology (Lin, 2010; Kolde et al., 2012), engineering (Gaertner, 2019), and machine learning (Prati, 2012; Sarkar et al., 2014). For instance, in machine learning, ranking aggregation is used to combine rankings produced by different machine learning algorithms into a final ranking to make the outcome more robust.

In social choice theory, the problem of ranking aggregation is formalized by *social preference functions (SPFs)*, which compute a non-empty set of strict winning rankings based on the voters' preferences.[1] In particular, the literature mainly focuses on two types of SPFs: positional scoring rules (e.g., Smith, 1973; Young, 1975; Myerson, 1995) and Kemeny's rule (e.g., Young and Levenglick, 1978; Barthelemy and Monjardet, 1981; Can and Storcken, 2013). In a positional scoring rule, each voter assigns a fixed score to each alternative depending on its position in the voter's preference relation and the output rankings arrange the alternatives in decreasing order of their total scores. An example of a positional scoring rule is Borda's rule, where voters give $m-1$ points to their favorite alternative, $m-2$ points to their second-best alternative, and so on. By contrast, in Kemeny's rule, voters give points to pairs of alternatives: a voter gives 1 point to a pair (a, b) if he prefers a to b and -1 point if he prefers b to a. Kemeny's rule then views each ranking $\rhd$ as a set of ordered pairs and the score of a ranking is the sum of scores assigned to the pairs $(a, b) \in \rhd$. Finally, the chosen rankings are those with maximal scores.

Intuitively, Kemeny's rule and scoring rules have a rather diametrical behavior. On the one hand, scoring rules are closely related to preference intensities: for example, Borda's rule assumes that a voter's preference between his first-ranked and last-ranked alternatives is much stronger than his preference between his first-ranked and second-ranked alternatives. This interpretation follows by simply considering the scores assigned by Borda's rule. On the other hand, Kemeny's rule completely ignores such preference intensities as each voter gives 1 point to a pair of alternatives (a, b) if he prefers a to b and -1 if he prefers b to a, independently of how close a and b are in the voter's preference relation. Indeed, Kemeny's rule can be seen as a natural extension of the majority rule (given the choice between two alternatives, pick the one that is preferred by a majority of the voters) to ranking aggregation. This has even been formally observed by Young and Levenglick (1978) who have characterized Kemeny's rule based on a variant

[1] Another approach to formalize ranking aggregation are social welfare functions (SWFs), which return a single weak ranking over the alternatives (see, e.g., Smith, 1973; Campbell and Kelly, 2002). However, every SWF can be turned into an SPF by returning all strict rankings that can be derived from the weak ranking by tie-breaking. By contrast, not every SPF can be turned into an SWF.

of Condorcet-consistency for rankings. By contrast, positional scoring rules inherently conflict with the concept of majority decisions (Fishburn and Gehrlein, 1976).

Despite these intuitive differences, positional scoring rules and Kemeny's rule are quite similar from an axiomatic perspective. This is, e.g., demonstrated by the characterization of positional scoring rules by Smith (1973) and the characterization of Kemeny's rule by Young (1988): both of these results are driven by mild fairness axioms and a consistency condition for variable electorates. In this paper, we thus aim to unify these two classical results by giving a joint axiomatic foundation of positional scoring rules and Kemeny's rule.

Our contribution. To derive the joint axiomatic basis of positional scoring rules and Kemeny's rule, we will introduce and fully axiomatize a new class of SPFs called bivariate scoring rules. For these rules, voters assign—just as for Kemeny's rule—scores to each pair of alternatives (a, b), but the exact score depends—analogous to positional scoring rules—on the positions of a and b in the voters' preference relations. The score of an output ranking $\rhd$ are then the sum of the scores assigned to pairs (a, b) with $(a, b) \in \rhd$ and the rule chooses the rankings with maximal total scores. In particular, all positional scoring rules and Kemeny's rule are bivariate scoring rules, so this class allows for a joint treatment of these SPFs. Moreover, bivariate scoring rules can be seen as variants of Kemeny's rule that incorporate preference intensities as it is, for example, possible to assign scores depending on the distance between alternatives. Such generalizations of Kemeny's rule have attracted significant attention (e.g., Cook and Kress, 1986; Kumar and Vassilvitskii, 2010; Can, 2014; Plaia et al., 2019) as it is a common observation that, e.g., the first positions in a input ranking are more important than the later ones.

For our characterization of bivariate scoring rules, we rely (aside from mild standard axioms) on consistency conditions for variable electorates and variable agendas of alternatives. In more detail, we use the prominent notion of reinforcement as variable electorate condition, which requires that if an SPF chooses some rankings for two disjoint elections, then precisely these common winning rankings are chosen in a combined election. Variants of this axiom feature in numerous prominent characterizations (e.g., Smith, 1973; Fishburn, 1978; Young and Levenglick, 1978; Brandl et al., 2016; Lackner and Skowron, 2021). For consistency with respect to variable agendas, we use a new axiom called local agenda consistency, which can be seen as a variant of Young's pairwise consistency (1988) or Sen's contraction consistency (1971; 1977). Intuitively, local agenda consistency requires that if a ranking $\rhd$ is chosen for a feasible set Y, then it holds for every set of *consecutive* alternatives X in $\rhd$ that $\rhd$ restricted to X is a winning ranking for X. If this were not the case, there would be a better ranking $\rhd'$ for the set X and we could improve the ranking $\rhd$ by reordering the alternatives in X according to $\rhd'$ because these alternatives appear consecutively in $\rhd$. As our main result, we then show that *an SPF is a bivariate scoring rule if and only if it satisfies reinforcement, local agenda consistency, anonymity, neutrality, continuity, and faithfulness (Theorem 1).*

Based on this result, we furthermore infer variants of the prominent characterizations of positional scoring rules by Smith (1973) and of Kemeny's rule by Young (1988), which demonstrates that our characterization of bivariate scoring rules indeed unifies these two independent lines of research. In more detail, for our characterization of positional scoring rules, we strengthen local agenda consistency to agenda consistency. This axiom requires that the winning rankings for a feasible set X can be derived from those of a feasible set Y with $X \subseteq Y$ by simply restricting the winning rankings for Y to the alternatives in X. For example, agenda consistency implies that if the ranking $a \succ b \succ c \succ d$ is uniquely chosen for the set $Y = \{a, b, c, d\}$, then $a \succ c \succ d$ is uniquely chosen for the set $X = \{a, c, d\}$. We then show that *an SPF is a positional scoring rule if and only if it satisfies reinforcement, agenda consistency, anonymity, neutrality, continuity, and faithfulness (Corollary 1).* This result is in the spirit of the prominent characterization of positional scoring rules by Smith (1973) because the conjunction of agenda consistency and reinforcement has similar consequences as separability, the main axiom of Smith (see Remark 5 for details). However, while Smith (1973) proves his result for social welfare functions, our result concerns the more general setting of SPFs.

As the second corollary, we characterize Kemeny's rule as the only bivariate scoring rule that satisfies independence of infeasible alternatives.[2] This axiom requires that the chosen rankings only depend on the alternatives in the feasible set (see, e.g., Arrow (1951) or Sen (2017, Chapter 3*)). In more detail, we prove that *an SPF is Kemeny's rule if and only if it satisfies reinforcement, local agenda consistency, independence of infeasible alternatives, anonymity, neutrality, continuity, and faithfulness (Corollary 2).* This result is similar to the characterization of Kemeny's rule by Young (1988) who uses mostly the same axioms but another variable agenda condition called pairwise consistency (we refer to Remark 6 for details).[3] Finally, we note that Corollaries 1 and 2 emphasize the differences between positional scoring rules and Kemeny's rule: while the former SPFs satisfy stronger agenda consistency conditions, Kemeny's rule is additionally independent of infeasible alternatives.

Related Work. Ranking aggregation is one of the most classic problems in social choice theory and it is therefore not surprising that various SPFs have been suggested in the literature, with the most prominent ones including Kemeny's rule (Kemeny, 1959), positional scoring rules (Smith, 1973), various types of runoff scoring rules (Smith, 1973; Boehmer et al., 2023), Slater's rule (Slater, 1961), and the ranked pairs method (Tideman, 1987). This paper will focus on positional scoring rules and Kemeny's rule; for an overview of other rules, we refer to Arrow et al. (2002) and Brandt et al. (2016).

Both Kemeny's rule and positional scoring rules have attracted significant attention and multiple characterizations are known for both types of rules. In more detail, Young

[2] In light of Arrow's impossibility theorem, it may sound surprising that Kemeny's rule (or any other attractive SPF) satisfies independence of infeasible alternatives. This claim holds because we use, in contrast to Arrow's impossibility theorem, a variable agenda framework. In this case, independence of infeasible alternatives does not impose any consistency conditions between different feasible sets and is therefore easy to satisfy. Indeed, when using a variable agenda framework, it is well-known that independence of infeasible alternatives only becomes prohibitive when combined with agenda consistency notions (see, e.g., Bordes, 1976; Sen, 1977).

[3] Young (1988) calls this concept first local stability and introduces the term pairwise consistency later on (Young, 1994, Theorem 6). Furthermore, pairwise consistency is sometimes called local independence of irrelevant alternatives (e.g., Young, 1995; Boehmer et al., 2023). We choose to call this condition pairwise consistency to avoid confusion with our other axioms.

and Levenglick (1978), Young (1988), and Can and Storcken (2013) prove characterizations of Kemeny's rule. For instance, Young and Levenglick (1978) characterize Kemeny's rule based on neutrality, reinforcement, and a non-standard notion of Condorcet-consistency (see Remark 7). Moreover, Bossert and Sprumont (2014) study Kemeny's rule with respect to manipulability and several authors geometrically compare Kemeny's rule to other SPFs (Saari and Merlin, 2000; Ratliff, 2001; Klamler, 2004), all of whom conclude that Kemeny's rule performs outstandingly from an axiomatic perspective. On the negative side, it is computationally intractable to compute the winning rankings for Kemeny's rule (Bartholdi, III et al., 1989; Hemaspaandra et al., 2005).

The picture for positional scoring rules is similar as there is a multitude of characterizations of this class (Smith, 1973; Young, 1974b, 1975; Myerson, 1995) and also of specific positional scoring rules such as Borda's rule (Young, 1974a; Nitzan and Rubinstein, 1981) or the plurality rule (Richelson, 1978; Ching, 1996); we refer to Chebotarev and Shamis (1998) for a more detailed overview of this extensive stream of research. Remarkably, all these characterizations crucially rely on variants of reinforcement as positional scoring rules satisfy this condition—in contrast to Kemeny's rule—also as social choice functions (which return a set of winning alternatives rather than a set of winning rankings). Moreover, there is also a vast number of results on specific aspects of scoring rules, such as their manipulability (Favardin and Lepelley, 2006; Pritchard and Wilson, 2007), their welfare guarantees (Pivato, 2016; Skowron and Elkind, 2017), or their probability to elect the Condorcet winner (Gehrlein, 1982; Cervone et al., 2005).

In this paper, we will derive a joint characterization of Kemeny's rule and positional scoring rules, thus unifying all of the above results. We note that there are already several classes of voting rules that could be used for this purpose, such as the mean neat rules of Zwicker (2008), the simple scoring ranking rules of Conitzer et al. (2009), or the distance-based voting rules studied in the context of distance rationalizability (Elkind and Slinko, 2016). However, these classes are too general to allow for an appealing characterization. For an example, consider the simple scoring ranking rules by Conitzer et al. (2009), which are defined by a scoring function $s(\succ, \rhd)$ that states how many points a voter with preference relation $\succ$ assigns to each output ranking $\rhd$ and choose the rankings with the maximal total score. While this class is easy to understand, it is extremely general and thus very challenging to axiomatize. By contrast, we derive an appealing characterization by focusing on the smaller class of bivariate scoring rules.

Finally, our results are conceptually related to numerous prominent results in social choice theory because variants of reinforcement and (local) agenda consistency have frequently been studied. For instance, reinforcement has been used to characterize positional scoring rules in the context of single winner elections (Young, 1975; Myerson, 1995; Pivato, 2013), approval voting (Fishburn, 1978; Brandl and Peters, 2022), and committee voting (Skowron et al., 2019; Lackner and Skowron, 2021; Dong and Lederer, 2024). Moreover, Brandl et al. (2016) use axioms similar to reinforcement and agenda consistency to characterize a randomized voting rule called maximal lotteries. The study of agenda consistency notions in social choice theory goes back to Sen (1971, 1977) and has since then attracted significant attention (e.g., Tideman, 1987; Laffond et al., 1996; Laslier, 1997). Hence, we employ classical ideas to derive new characterizations.

2. The Model

Let $\mathbb{N} = \{1, 2, 3, \dots\}$ denote an infinite set of voters and $A = \{a_1, \dots, a_m\}$ denote a set of m alternatives. In this paper, we employ both a variable agenda and a variable population framework. To this end, we define $\mathcal{F}(X) = \{Y \subseteq X : Y \text{ is non-empty and finite}\}$ as the set of all finite and non-empty subsets of a given set X. In particular, we interpret $\mathcal{F}(\mathbb{N})$ as the set of possible electorates, whereas $\mathbb{N}$ is the set of all possible voters. Similarly, $\mathcal{F}(A)$ is the set of all possible agendas (or feasible sets) of alternatives. Intuitively, a set $X \in \mathcal{F}(A)$ contains the alternatives that need to be ranked.

Given an electorate $N \in \mathcal{F}(\mathbb{N})$, each voter $i \in N$ is assumed to report a *preference relation* $\succ_i$ on A, which is formally a strict total order on A. We call preference relations also *rankings* and typically write these objects as comma-separated lists; e.g., $\succ_i = (a, b, c)$ means that voter i prefers a to b to c. We will omit the brackets around preference relations whenever this helps readability. The set of all rankings over a set of alternatives X is $\mathcal{R}(X)$. Moreover, $\succ|_X$ denotes the restriction of a ranking $\succ \in \mathcal{R}(Y)$ to the alternatives in $X \subseteq Y$, i.e., $\succ|_X = \succ \cap X^2$. We call a set of alternatives X *consecutive* in a ranking $\succ \in \mathcal{R}(Y)$ if $x \succ z$ if and only if $y \succ z$ for all $x, y \in X$, $z \in Y \setminus X$. Finally, we define $C(X, Y) = \{\succ \in \mathcal{R}(Y) : X \text{ is consecutive in } \succ\}$ as the set of rankings on Y in which the alternatives in X appear consecutively.

A *preference profile* R for an electorate $N \in \mathcal{F}(\mathbb{N})$ is a mapping from N to $\mathcal{R}(A)$, i.e., it contains the preference relation of every voter $i \in N$. Conversely, N_R denotes the set of voters that report a preference relation in R. We note that, even though we allow for a variable agenda, preference profiles are always defined on all alternatives. When writing preference profiles, the number before a preference relation indicates how many voters report a given preference relation; for instance, $3\colon a, b, c$ means that three voters prefer a to b to c. Next, we define $\mathcal{R}^*$ as the set of all possible preference profiles: $\mathcal{R}^* = \bigcup_{N \in \mathcal{F}(\mathbb{N})} \mathcal{R}(A)^N$. Furthermore, we call two profiles $R, R' \in \mathcal{R}^*$ disjoint if $N_R \cap N_{R'} = \emptyset$ and define the profile $R'' = R + R'$ for two disjoint profiles R, R' by $N_{R''} = N_R \cup N_{R'}$, $\succ''_i = \succ_i$ for all $i \in N_R$, and $\succ''_i = \succ'_i$ for all $i \in N_{R'}$.

In this paper, we aim to study *social preference functions (SPFs)* which choose a non-empty set of winning rankings $W \subseteq \mathcal{R}(X)$ for each preference profile $R \in \mathcal{R}^*$ and each feasible set of alternatives $X \in \mathcal{F}(A)$. More formally, an SPF is a function f of the type $\mathcal{R}^* \times \mathcal{F}(A) \to \bigcup_{X \in \mathcal{F}(A)} \mathcal{F}(\mathcal{R}(X))$ such that $f(R, X) \subseteq \mathcal{R}(X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$. While SPFs are also defined for feasible sets X of size 1, we will ignore this case as there is only a single ranking if $|X| = 1$. The same model has implicitly been considered by, e.g., Young (1988) and Boehmer et al. (2023). Finally, to clearly distinguish between the preference relations of voters and the winning rankings chosen by an SPF, we use $\succ$ for the former and $\rhd$ for the latter.

2.1. Bivariate Scoring Rules

In this paper, we focus on three types of SPFs: positional scoring rules, Kemeny's rule, and our new class of bivariate scoring rules. An example illustrating these different types of SPFs is shown in Figure 1.

Positional Scoring Rules. We first introduce positional scoring rules and define for this the rank of an alternative x in a preference relation $\succ_i$ as $r(\succ_i, x) = 1 + |\{y \in A \setminus \{x\}\colon y \succ_i x\}|$. Less formally, $r(\succ_i, x) = k$ means that x is the k-th best alternative of voter i. Each positional scoring rule is defined by a positional scoring function $s : \{1, \dots, m\} \to \mathbb{R}$ which is non-increasing and non-constant, i.e., it holds that $s(1) \geq s(2) \geq \dots \geq s(m)$ and $s(1) > s(m)$.[4] Intuitively, $s(\ell)$ states how many points a voter assigns to his ℓ-th best alternative. Next, we define the total score of an alternative x in a profile R as $\hat{s}(R, x) = \sum_{i \in N_R} s(r(\succ_i, x))$. Finally, an SPF f is a *positional scoring rule* if there is a positional scoring function s such that $f(R, X) = \{\rhd \in \mathcal{R}(X)\colon \forall x, y \in X\colon x \rhd y \implies \hat{s}(R, x) \geq \hat{s}(R, y)\}$ for all profiles $R \in \mathcal{R}^*$ and all feasible sets $X \in \mathcal{F}(A)$. We note that positional scoring rules are defined in a "broad" sense: we always compute the scores with respect to the full profile R on A, even if some alternatives are not in the feasible set. Common examples of positional scoring rules are Borda's rule (defined by $s(x) = m - x$) or the plurality rule (defined by $s(1) = 1$ and $s(x) = 0$ for $x > 1$).

Kemeny's Rule. Next, we turn to Kemeny's rule which interprets rankings $\rhd \in \mathcal{R}(X)$ as a set of ordered pairs of alternatives, i.e., $\rhd = \{(x, y) \in X^2\colon x \rhd y\}$. Formally, Kemeny's rule is also defined by a scoring function s which takes two alternatives x, y and a preference relation $\succ_i$ as input: $s(x, y, \succ_i) = 1$ if $x \succ_i y$ and $s(x, y, \succ_i) = -1$ if $y \succ_i x$. The Kemeny score of a ranking $\rhd \in \mathcal{R}(X)$ in a profile R is then defined as $\hat{s}_{Kemeny}(R, \rhd) = \sum_{(x,y) \in \rhd} \sum_{i \in N_R} s(x, y, \succ_i)$ and Kemeny's rule chooses the rankings with maximal Kemeny score, i.e., $f_{Kemeny}(R, X) = \{\rhd \in \mathcal{R}(X)\colon \forall \rhd' \in \mathcal{R}(X)\colon \hat{s}_{Kemeny}(R, \rhd) \geq \hat{s}_{Kemeny}(R, \rhd')\}$ for all profiles $R \in \mathcal{R}^*$ and feasible sets of alternatives $X \in \mathcal{F}(A)$. This definition is equivalent to choosing the rankings that minimize the total swap distance to the voters' preference relations. We note that Kemeny's rule reduces to the majority rule for agendas of size 2: $f_{Kemeny}(R, \{x, y\}) = \{(x, y)\}$ if a strict majority of voters prefers x to y, $f_{Kemeny}(R, \{x, y\}) = \{(y, x)\}$ if a strict majority prefers y to x, and $f_{Kemeny}(R, \{x, y\}) = \{(x, y), (y, x)\}$ otherwise.

Bivariate Scoring Rules. Finally, we introduce the class of bivariate scoring rules. To this end, we observe that Kemeny's rule can also be defined based on a scoring function s that only takes the ranks of two alternatives as input: $s(r(\succ_i, x), r(\succ_i, y)) = 1$ if $r(\succ_i, x) < r(\succ_i, y)$ and $s(r(\succ_i, x), r(\succ_i, y)) = -1$ if $r(\succ_i, x) > r(\succ_i, y)$. The idea for bivariate scoring rules is to introduce weights in this function. We hence define bivariate scoring functions as functions of the type $s : \{1, \dots, m\} \times \{1, \dots, m\} \to \mathbb{R}$ that satisfy $s(\ell, k) \geq 0$ and $s(\ell, k) = -s(k, \ell)$ for all $\ell, k \in \{1, \dots, m\}$ with $\ell \leq k$ and $s(\ell, k) > 0$ for some $\ell, k \in \{1, \dots, m\}$. Intuitively, a bivariate scoring function quantifies how important it is for a voter that his ℓ-th best alternative is ranked ahead of his k-th

[4] Positional scoring rules can be defined without these assumptions. However, if a positional scoring function s fails to be non-increasing, the corresponding positional scoring rule violates several desirable properties. Thus, numerous authors assume that the score vector is non-decreasing (e.g., Gehrlein, 1982; Pritchard and Wilson, 2007; Favardin and Lepelley, 2006; Pivato, 2016; Skowron and Elkind, 2017). The condition that s is non-constant only entails that the corresponding positional scoring rule does not always choose every ranking.

$$
R: \begin{array}{ll} 3\colon & a,b,c,d \\ 2\colon & b,c,d,a \\ 2\colon & c,d,a,b \end{array}
\qquad
\begin{array}{c} f_{Borda}(R,\{a,b,c\}) = \{(c,b,a)\} \\ f_{Kemeny}(R,\{a,b,c\} = \{(a,b,c)\} \\ f_d(R,\{a,b,c\}) = \{(c,b,a)\} \end{array}
$$

Figure 1: Examples of SPFs. The left side depicts the profile R which consists of 7 voters and 4 alternatives. On the right side, we display the rankings chosen by Borda's rule, Kemeny's rule, and the bivariate scoring rule f_d defined by $s(\ell, k) = -(\ell - k)^3$ for the feasible set $\{a, b, c\}$. It can be verified that Borda's rule assigns 11 points to a, 12 points to b, and 13 points to c (and 6 points to d), so it chooses the ranking c, b, a. On the other hand, 5 voters prefer a to b, 5 voters prefer b to c, and 4 voters prefer c to a. Consequently, the ranking a, b, c maximizes the Kemeny score with a value of $5-2+5-2+4-3 = 7$ and is therefore selected by Kemeny's rule. Finally, for computing f_d, we consider every voter and each pair of alternatives to compute the scores. For instance, each voter with the preference relation b, c, d, a gives 1 point to the pair (b, c), 4 points to the pair (c, a), and 27 points to the pair (b, a) (or conversely -1 point to (c, b), -4 points to (a, c), and -27 points to (a, b)). By continuing these computations, one can verify that c, b, a maximizes the total score as the pairs (c, b) and (b, a) both get $5 \cdot (-1) + 2 \cdot 27 = 49$ points.

best alternative in the output ranking. We thus define the score of a pair of alternatives (x, y) in a profile R as $\hat{s}(R, (x, y)) = \sum_{i \in N_R} s(r(\succ_i, x), r(\succ_i, y))$ and the score of a ranking $\rhd \in \mathcal{R}(X)$ as $\hat{s}(R, \rhd) = \sum_{(x,y) \in \rhd} \hat{s}(R, (x, y))$. Finally, an SPF f is a *bivariate scoring rule* if there is a bivariate scoring function s such that $f(R, X) = \{\rhd \in \mathcal{R}(X)\colon \forall \rhd' \in \mathcal{R}(X)\colon \hat{s}(R, \rhd) \geq \hat{s}(R, \rhd')\}$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$. More intuitively, bivariate scoring rules allow the election designer to value each comparison between two alternatives x, y in a voters' preference relation differently depending on the positions of x and y. Thus, these SPFs can be seen as weighted variants of Kemeny's rule. Finally, we note that the assumption that $s(\ell, k) = -s(k, \ell)$ is without loss of generality because the resulting bivariate scoring rule is independent of additive shifts for each pair of indices.

The class of bivariate scoring rules contains a variety of interesting SPFs, with Kemeny's rule constituting the most apparent example. Moreover, bivariate scoring rules are a very flexible concept, and one can, for instance, also define "distance-based" rules whose bivariate scoring function $s(\ell, k)$ only depends on $\ell - k$. For example, the rule f_d in Figure 1, which is defined by $s(\ell, k) = -(\ell - k)^3$, is such a rule. Finally, we will show that all positional scoring rules are bivariate scoring rules.

Proposition 1. *Every positional scoring rule is a bivariate scoring rule.*

Proof. Consider an arbitrary positional scoring rule f and let s denote its positional scoring function. We define a bivariate scoring function s' by $s'(\ell, k) = s(\ell) - s(k)$ for all $\ell, k \in \{1, \dots, m\}$ and let f' denote the corresponding bivariate scoring rule. First, we note that s' is indeed a bivariate scoring function since $s(1) \geq s(2) \geq \dots \geq s(m)$

and $s(1) > s(m)$, so $s'(\ell, k) \geq 0$ if $\ell \leq k$ and $s'(1, m) > 0$. Our goal is to show that $f(R, X) = f'(R, X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$. To this end, we note that $\hat{s}'(R, (x, y)) = \sum_{i \in N_R} s(r(\succ_i, x)) - s(r(\succ_i, y)) = \hat{s}(R, x) - \hat{s}(R, y)$ for all profiles $R \in \mathcal{R}^*$ and alternatives $x, y \in A$. Hence, it holds for every ranking $\rhd \in f'(R, X)$ and all alternatives $x, y \in X$ that $\hat{s}(R, x) \geq \hat{s}(R, y)$ if x is placed directly above y in $\rhd$. Otherwise, the ranking $\rhd'$ derived from $\rhd$ by swapping x and y has a higher score than $\rhd$ since $\hat{s}'(R, \rhd') = \hat{s}'(R, \rhd) - \hat{s}'(R, (x, y)) + \hat{s}'(R, (y, x)) = \hat{s}'(R, \rhd) + 2\hat{s}(R, y) - 2\hat{s}(R, x) > \hat{s}'(R, \rhd)$. This contradicts the assumption that $\rhd \in f'(R, X)$, so it holds for all rankings $\rhd \in f'(R, X)$ that $\hat{s}(R, x) \geq \hat{s}(R, y)$ if $x \rhd y$ and x, y are consecutive in $\rhd$. Even more, if there is a non-consecutive pair of alternatives $x, y \in X$ in $\rhd$ such that $\hat{s}(R, x) < \hat{s}(R, y)$ and $x \rhd y$, we can also find a consecutive pair that satisfies these conditions. Hence, we infer that $f'(R, X) \subseteq f(R, X)$. For the converse direction, we note that if $\hat{s}(R, x) = \hat{s}(R, y)$ for two alternatives $x, y \in X$ that are consecutive in a ranking $\rhd$, then $\hat{s}'(R, \rhd) = \hat{s}'(R, \rhd')$ for the ranking $\rhd'$ derived from $\rhd$ by swapping x and y. Hence, $\rhd \in f'(R, X)$ also implies that $\rhd' \in f'(R, X)$ for these rankings. Since we already know that $f'(R, X) \subseteq f(R, X)$ and the rankings in $f(R, X)$ only differ by reordering alternatives $x, y \in X$ with $\hat{s}(R, x) = \hat{s}(R, y)$, it follows that $f(R, X) = f'(R, X)$. This shows that f is indeed the bivariate scoring rule defined by s'. $\square$

2.2. Basic Axioms

We next introduce five standard axioms which will form the basis of our analysis. We note that these axioms can be used to characterize scoring rules in various contexts (Young, 1975; Myerson, 1995; Skowron et al., 2019; Lackner and Skowron, 2022). By contrast, this is not the case for SPFs as, e.g., scoring runoff rules also satisfy all axioms listed in this section.

Anonymity. Anonymity is a mild fairness condition that postulates that voters are treated equally. In more detail, an SPF f is *anonymous* if $f(R, X) = f(\pi(R), X)$ for all profiles $R \in \mathcal{R}^*$, feasible sets $X \in \mathcal{F}(A)$, and permutations $\pi : \mathbb{N} \to \mathbb{N}$. Here, $R' = \pi(R)$ denotes the profile such that $N_{R'} = \{\pi(i) \colon i \in N_R\}$ and $\succ'_{\pi(i)} = \succ_i$ for all $i \in N_R$.

Neutrality. Similar to anonymity, neutrality is a standard fairness condition for alternatives. Formally, an SPF f is *neutral* if $f(\tau(R), \tau(X)) = \{\tau(\rhd) \colon \rhd \in f(R, X)\}$ for all profiles $R \in \mathcal{R}^*$, feasible sets $X \in \mathcal{F}(A)$, and permutations $\tau : A \to A$. Here, $\rhd' = \tau(\rhd)$ is the ranking defined by $\tau(x) \rhd' \tau(y)$ if and only if $x \rhd y$ for all $x, y \in A$, and $R' = \tau(R)$ is the profile with $N_{R'} = N_R$, and $\succ'_i = \tau(\succ_i)$ for all $i \in N_R$.

Faithfulness. Faithfulness requires that SPFs respect the voters' preferences if there is only one voter and two alternatives in the feasible set. In more detail, an SPF f is *faithful* if for all alternatives $x, y \in A$, it holds that $(x, y) \in f(\succ_i, \{x, y\})$ for all rankings $\succ_i \in \mathcal{R}(A)$ with $x \succ_i y$ and $f(\succ_i, \{x, y\}) = \{(x, y)\}$ for some ranking $\succ_i \in \mathcal{R}(A)$ with $x \succ_i y$. This condition is, for instance, weaker than Pareto-optimality (which requires that if $x \succ_i y$ for all voters i in some profile R, then $x \rhd y$ for all $\rhd \in f(R, X)$).

Continuity. Continuity, also known as the overwhelming majority property (Myerson, 1995) or the Archimedean property (Smith, 1973), roughly states that large groups of voters can ensure that some of their desired outcomes are chosen. More formally, an SPF f is called *continuous* if for all profiles $R, R' \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$, there is $\lambda \in \mathbb{N}$ such that $f(\lambda R + R', X) \subseteq f(R, X)$. Here, the profile λR consists of λ copies of R; the names of the voters in λR will not matter due to anonymity. We note that all axioms defined so far are very mild and satisfied by all commonly studied SPFs.

Reinforcement. As the last axiom in this section, we introduce reinforcement. Intuitively, this axiom requires that if some rankings are chosen for two disjoint elections, then precisely the common winning rankings are chosen in a joint election. We hence call an SPF f *reinforcing* if $f(R, X) \cap f(R', X) \neq \emptyset$ implies that $f(R + R', X) = f(R, X) \cap f(R', X)$ for all disjoint profiles $R, R' \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$. The idea here is that a winning ranking is as good as any other winning ranking and better than any unchosen ranking. Thus, if some rankings are chosen for two disjoint elections, these are the best rankings for the combined election.

2.3. Agenda consistency

As the last ingredient for our results, we will introduce consistency conditions for variable agendas. Roughly, these axioms describe how the outcome should change if we modify the feasible set. To make this more clear, let us revisit the profile R in Figure 1 where $f_{Borda}(R, \{a, b, c\}) = \{(c, b, a)\}$. The question we now ask is how the winning ranking changes when, e.g., alternative b is deleted from the feasible set. Maybe the most straightforward idea is that we should choose the ranking (c, a)—we simply remove the now unavailable alternative b from the ranking to derive the winning ranking for the smaller feasible set. We formalize this idea with *agenda consistency* which postulates of an SPF f that $f(R, X) = \{\rhd|_X : \rhd \in f(R, Y)\}$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X, Y \in \mathcal{F}(A)$ with $X \subseteq Y$. Note that the general idea of this axiom coincides with known rationality conditions such as Sen's contraction and expansion consistency for (social) choice functions (Sen, 1971, 1977). Even though agenda consistency is a restrictive axiom, all positional scoring rules satisfy this condition as these rules always order the feasible alternatives with respect to the scores computed for the complete profile.

Unfortunately, it turns out that Kemeny's rule fails agenda consistency. This can be shown by considering the profile R in Figure 1: it holds that $f_{Kemeny}(R, \{a, c\}) = \{(c, a)\}$ but $f_{Kemeny}(R, \{a, b, c\}) = \{(a, b, c)\}$. At a first glance, this may seem like a flaw of Kemeny's rule because the ordering over a and c depends on the availability of b. However, the problem runs much deeper: agenda consistency entails a high degree of transitivity for agendas of size 2. In more detail, since $f_{Kemeny}(R, \{a, b\}) = \{(a, b)\}$, $f_{Kemeny}(R, \{b, c\}) = \{(b, c)\}$, and $f_{Kemeny}(R, \{a, c\}) = \{(c, a)\}$, f_{Kemeny} fails agenda consistency *regardless of which rankings it chooses for* $\{a, b, c\}$. Or, put differently, agenda consistency requires a form of transitivity for the choice on agendas of size 2: if $f(R, \{a, b\}) = \{(a, b)\}$ and $f(R, \{b, c\}) = \{(b, c)\}$, then $f(R, \{a, c\}) = \{(a, c)\}$. In concordance with a large stream of research (e.g., May, 1954; Fishburn, 1970; Bar-Hillel and Margalit, 1988; Anand, 2009), we view this transitivity requirement of agenda consistency as unduly restrictive because such transitivity notions have empirically been shown to be unrealistic and often lead to negative theoretical results.

We therefore consider a weakening of agenda consistency, which is based on the idea that a winning ranking $\rhd$ for a large set Y should only inherit to a subset X if the alternatives in X appear consecutively in $\rhd$. The reason for this is that if $\rhd|_X$ is not chosen for the set X, there is a better ranking for these alternatives. Moreover, we can reorder the alternatives in X in the ranking $\rhd$ without affecting the alternatives in $Y \setminus X$, so we can intuitively also improve this ranking. To formalize this idea, we recall that $C(X,Y)$ denotes the set of rankings $\rhd \in \mathcal{R}(Y)$ in which the alternatives in X appear consecutively. Then, *local agenda consistency* requires that $f(R,X) \supseteq \{\rhd|_X \colon \rhd \in f(R,Y) \cap C(X,Y)\}$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X, Y \in \mathcal{F}(A)$ with $X \subseteq Y$. This is a rather mild condition and we will later show that all bivariate scoring rules satisfy it.

Finally, we will introduce Arrow's independence of infeasible alternatives (Arrow, 1951), which demands that the winning rankings should only depend on the available alternatives. More formally, an SPF f satisfies *independence of infeasible alternatives* if $f(R,X) = f(R',X)$ for all profiles $R, R' \in \mathcal{R}^*$ and all feasible sets $X \in \mathcal{F}(A)$ such that $N_R = N_{R'}$ and $\succ_i|_X = \succ'_i|_X$ for all $i \in N_R$. In our context, independence of infeasible alternatives is easy to satisfy as it suffices to apply a voting rule to the input profile restricted to the feasible set. Moreover, this axiom is typically motivated by the fact that it allows to compute the outcome for a given feasible set only depending on the voters' preferences over this set of alternatives. By contrast, just as Sen's contraction consistency, our agenda consistency notions only relate the outcomes for different feasible sets but do not restrict the information that can be used to compute the winning rankings. As a consequence, (local) agenda consistency and independence of infeasible alternatives are logically independent: for instance, all "narrow" positional scoring rules, which first restrict the profile to the feasible alternatives and then compute the scores of the alternatives, satisfy independence of infeasible alternatives but fail local agenda consistency, and Corollary 2 shows that all bivariate scoring rules except of Kemeny's rule satisfy local agenda consistency but fail independence of infeasible alternatives.

3. Characterizations

We are now ready to state our results. In more detail, we show in Section 3.1 the characterization of bivariate scoring rules, and discuss in Sections 3.2 and 3.3 the characterizations of positional scoring rules and Kemeny's rule, respectively.

3.1. Bivariate Scoring Rules

In this section, we discuss our main result, the characterization of bivariate scoring rules, and its consequences. Since the proof of this result is rather involved, we only show in the main body that every bivariate scoring rule satisfies all given axioms and give a proof sketch for the inverse direction. The full proof can be found in the appendix.

Theorem 1. *An SPF is a bivariate scoring rule if and only if it satisfies anonymity, neutrality, continuity, faithfulness, reinforcement, and local agenda consistency.*

Proof Sketch. The proof consists of two claims: we need to show that bivariate scoring rules satisfy all given axioms and that they are the only SPFs that do so.

($\Longrightarrow$) We first show that every bivariate scoring rule satisfies all considered properties. Hence, let f denote a bivariate scoring rule and let s denote its corresponding bivariate scoring function. We first note that f is anonymous and neutral as s does not depend on the names of voters or alternatives. Furthermore, f is faithful because $s(\ell, k) \geq 0$ for all $\ell, k \in \{1, \dots, m\}$ with $\ell \leq k$ and $s(\ell, k) > 0$ for some $\ell, k \in \{1, \dots, m\}$. Next, f satisfies continuity because it maximize scores: it holds for all profiles $R \in \mathcal{R}^*$, feasible sets $X \in \mathcal{F}(A)$, and rankings $\rhd \in f(R, X)$ and $\rhd' \in \mathcal{R}(X) \setminus f(R, X)$ that $\hat{s}(R, \rhd) > \hat{s}(R, \rhd')$. Thus, there is for every other profile R' a $\lambda \in \mathbb{N}$ such that $\hat{s}(\lambda R + R', \rhd) > \hat{s}(\lambda R + R', \rhd')$ for all $\rhd \in f(R, X)$, $\rhd' \in \mathcal{R}(X) \setminus f(R, X)$, which entails that $f(\lambda R + R', X) \subseteq f(R, X)$.

Next, we will prove that f is reinforcing, for which we consider two disjoint profiles $R, R' \in \mathcal{R}^*$ and a feasible set $X \in \mathcal{F}(A)$ such that $f(R, X) \cap f(R', X) \neq \emptyset$. This means for all rankings $\rhd \in f(R, X) \cap f(R', X)$ and $\rhd' \in \mathcal{R}(X)$ that $\hat{s}(R, \rhd) \geq \hat{s}(R, \rhd')$ and $\hat{s}(R', \rhd) \geq \hat{s}(R', \rhd')$, so $\hat{s}(R + R', \rhd) \geq \hat{s}(R + R', \rhd')$ and $\rhd \in f(R + R', X)$. Moreover, if a ranking $\rhd'$ is not in $f(R, X)$ or in $f(R', X)$, then $\hat{s}(R, \rhd) > \hat{s}(R, \rhd')$ or $\hat{s}(R', \rhd) > \hat{s}(R', \rhd')$ for every $\rhd \in f(R, X) \cap f(R', X)$. This implies that $\hat{s}(R + R', \rhd) > \hat{s}(R + R', \rhd')$ and $\rhd' \notin f(R + R, X)$. Hence, $f(R + R', X) = f(R, X) \cap f(R', X)$ and f satisfies reinforcement.

Finally, we show that f satisfies local agenda consistency. To this end, consider two feasible sets $X, Y \in \mathcal{F}(A)$ with $X \subseteq Y$, a profile $R \in \mathcal{R}^*$, and a ranking $\rhd \in f(R, Y) \cap C(X, Y)$. The score of the ranking $\rhd$ in R is $\hat{s}(R, \rhd) = \hat{s}(R, \rhd|_X) + \sum_{(x,y) \in \rhd \setminus X^2} \hat{s}(R, (x, y))$. Now, if $\rhd|_X \notin f(R, X)$, there is a ranking $\rhd' \in \mathcal{R}(X)$ such that $\hat{s}(R, \rhd') > \hat{s}(R, \rhd|_X)$. Since the alternatives in X appear consecutively in $\rhd$, we can reorder these alternatives in $\rhd$ according to $\rhd'$ to derive another ranking $\rhd'' \in \mathcal{R}(Y)$ such that $\hat{s}(R, \rhd'') = \hat{s}(R, \rhd') + \sum_{(x,y) \in \rhd \setminus X^2} \hat{s}(R, (x, y)) > \hat{s}(R, \rhd|_X) + \sum_{(x,y) \in \rhd \setminus X^2} \hat{s}(R, (x, y)) = \hat{s}(R, \rhd)$. Hence, if $\rhd|_X \notin f(R, X)$, then $\rhd \notin f(R, Y)$, which contradicts our assumption and therefore shows that f satisfies local agenda consistency.

($\Longleftarrow$) Let f denote an SPF that satisfies anonymity, neutrality, continuity, faithfulness, reinforcement, and local agenda consistency. The main goal of our proof is to find a bivariate scoring function s that describes f. To this end, we will use a similar approach as presented by, e.g., Young (1975) or Young and Levenglick (1978), but we will need several new ideas to make our proof work. We will present only a high-level overview of our proof here and postpone all technical details to the appendix.

As a first step, we will show that the conjunction of our axioms implies that f is non-imposing. This means that for every feasible set $X \in \mathcal{F}(A)$ and every ranking $\rhd \in \mathcal{R}(X)$, there is a profile R such that $f(R, X) = \{\rhd\}$. Next, we will change the domain of f from preference profiles to a numerical space. To this end, we represent preference profiles R by vectors $v \in \mathbb{N}_0^{m!}$ which state how often each possible preference relation is reported. In particular, $v(R)$ denotes the vector corresponding to the profile R. Since f is anonymous, it is straightforward that there is a function g such that $f(R, X) = g(v(R), X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$. Moreover, it can be shown that g inherits all desirable properties of f. As the next step, we use reinforcement to extend the domain of g to $\mathbb{Q}^{m!} \times \mathcal{F}(A)$ while preserving the desirable properties of f. This leads to a new function $\hat{g}$ which still satisfies that $f(R, X) = \hat{g}(v(R), X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$. Hence, it suffices to describe the function $\hat{g}$ by a bivariate scoring function to show that f is a bivariate scoring rule.

To this end, we define the sets $D_\rhd = \{v \in \mathbb{Q}^{m!} \colon \rhd \in \hat{g}(v, X)\}$ for all feasible sets $X \in \mathcal{F}(A)$ and rankings $\rhd \in \mathcal{R}(X)$, i.e., $D_\rhd$ is the subdomain of $\mathbb{Q}^{m!}$ for which $\rhd \in \hat{g}(v, X)$. Moreover, we denote by $\bar{D}_\rhd$ the closure of $D_\rhd$ with respect to $\mathbb{R}^{m!}$ and observe that $\hat{g}(v, X) = \{\rhd \in \mathcal{R}(X) \colon v \in D_\rhd\} \subseteq \{\rhd \in \mathcal{R}(X) \colon v \in \bar{D}_\rhd\}$. So, our next intermediate goal is to describe the sets $\bar{D}_\rhd$. Here, it can be shown that the sets $\bar{D}_\rhd$ are convex as $\hat{g}$ inherits reinforcement of f, and that $\mathrm{int}\bar{D}_\rhd \cap \mathrm{int}\bar{D}_{\rhd'} = \emptyset$ for all feasible set $X \in \mathcal{F}(A)$ and rankings $\rhd, \rhd' \in \mathcal{R}(X)$. As a consequence, we can apply the separating hyperplane theorem for convex sets to find a non-zero vector $u^{\rhd,\rhd'} \in \mathbb{R}^{m!}$ such that $vu^{\rhd,\rhd'} \geq 0$ for all $v \in \bar{D}_\rhd$ and $vu^{\rhd,\rhd'} \leq 0$ for all $v \in \bar{D}_{\rhd'}$ (where vu denotes the standard scalar product). These vectors describe the sets $\bar{D}_\rhd$ as we prove that $\bar{D}_\rhd = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathcal{R}(X) \setminus \{\rhd\} \colon vu^{\rhd,\rhd'} \geq 0\}$ and we will therefore study the vectors $u^{\rhd,\rhd'}$ next.

To do so, we first restrict our attention to feasible sets of size 2 because there are only two possible rankings if $X = \{x, y\}$, namely (x, y) and (y, x). Consequently, $\bar{D}_{(x,y)} = \{v \in \mathbb{R}^{m!} \colon vu^{(x,y),(y,x)} \geq 0\}$, so it suffices to study a single vector $u^{(x,y),(y,x)}$. We therefore derive a bivariate scoring function s such that $u_k^{(x,y),(y,x)} = s(r(\succ^k, x), r(\succ^k, y)) - s(r(\succ^k, y), r(\succ^k, x))$ for all alternatives $x, y \in A$ and preference relations $\succ^k \in \mathcal{R}(A)$. Next, we turn to larger feasible sets and note for this that $\hat{g}$ inherits local agenda consistency from f. This implies that $\bar{D}_\rhd \subseteq \bar{D}_{(x,y)}$ if x and y appear consecutively in $\rhd$ and $x \rhd y$. As a consequence, it holds for all rankings $\rhd, \rhd' \in \mathcal{R}(X)$ with $\rhd \setminus \rhd' = \{(x, y)\}$ that the vector $u^{(x,y),(y,x)}$ separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$ (i.e., $vu^{(x,y),(y,x)} \geq 0$ if $v \in \bar{D}_\rhd$ and $vu^{(x,y),(y,x)} \leq 0$ if $v \in \bar{D}_{\rhd'}$) because it separates $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$. The reason for this is that $\rhd \setminus \rhd' = \{(x, y)\}$ implies that x and y appear consecutively in both $\rhd$ and $\rhd'$ but in different order.

Next, we aim to understand the vectors $u^{\rhd,\rhd'}$ for rankings $\rhd, \rhd'$ with $|\rhd \setminus \rhd'| > 1$. To do so, we fix a feasible set $X = \{a_1, \dots, a_\ell\}$ and introduce a neighbor relation on rankings in $\mathcal{R}(X)$. In more detail, we associate every ranking $\rhd \in \mathcal{R}(X)$ with the matrix $M^\rhd$ defined by $M^\rhd_{i,j} = 1$ if $a_i \rhd a_j$ and $M^\rhd_{i,j} = 0$ otherwise. Next, we define the set $\mathcal{M}$ as the convex hull of these matrices, and call two rankings $\rhd, \rhd'$ neighbors if they are neighboring extreme points in $\mathcal{M}$. This neighborhood relation has also been used by Young and Levenglick (1978) because there is a characterization of neighbors by Young (1978) (see also Gilmore and Hoffmann (1964)), which turns out helpful in the analysis. In particular, we show for all neighboring rankings $\rhd, \rhd' \in \mathcal{R}(X)$ that the vector $\sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$. We note that this step is the technically most involved one and includes, e.g., a complete specification of the linear (in)dependence of the vectors $u^{(x,y),(y,x)}$.

We continue by defining the set $\mathit{Neighbor}(\rhd)$ as the neighbors of a ranking $\rhd \in \mathcal{R}(X)$, and let $\bar{D}^N_\rhd = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathit{Neighbor}(\rhd) \colon vu^{\rhd,\rhd'} \geq 0\}$. It holds that $\hat{g}(v, X) \subseteq \{\rhd \in \mathcal{R}(X) \colon v \in \bar{D}^N_\rhd\}$ for all $v \in \mathbb{Q}^{m!}$ and $X \in \mathcal{F}(A)$ since $\bar{D}_\rhd \subseteq \bar{D}^N_\rhd$. Now, by the last step, we can choose $u^{\rhd,\rhd'} = \sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ for all $\rhd \in \mathcal{R}(X)$, $\rhd' \in \mathit{Neighbor}(\rhd)$.

Moreover, we already know that the vectors $u^{(x,y),(y,x)}$ can be described by a bivariate scoring function s. This implies for neighboring rankings $\rhd, \rhd' \in \mathcal{R}(X)$ that

$$\begin{aligned}\hat{s}(v, \rhd) - \hat{s}(v, \rhd') &= \sum_{(x,y) \in \rhd \setminus \rhd'} \hat{s}(v, (x, y)) - \sum_{(x,y) \in \rhd' \setminus \rhd} \hat{s}(v, (x, y)) \\ &= v \sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)},\end{aligned}$$

where $\hat{s}(v, (x, y)) = \sum_{k=1}^{m!} v_k s(r(\succ^k, x), r(\succ^k, y))$ and $\hat{s}(v, \rhd) = \sum_{(x,y) \in \rhd} \hat{s}(v, (x, y))$. Hence, we infer that $\bar{D}^N_{\rhd} = \{v \in \mathbb{R}^{m!} : \forall \rhd' \in \mathit{Neighbor}(\rhd) : \hat{s}(v, \rhd) \geq \hat{s}(v, \rhd')\}$.

Finally, we note that $\{v \in \mathbb{R}^{m!} : \forall \rhd' \in \mathit{Neighbor}(\rhd) : \hat{s}(v, \rhd) \geq \hat{s}(v, \rhd')\} = \{v \in \mathbb{R}^{m!} : \forall \rhd' \in \mathcal{R}(X) : \hat{s}(v, \rhd) \geq \hat{s}(v, \rhd')\}$. The reason for this is that the second set can be seen as a linear optimization problem: given a vector $v \in \mathbb{R}^{m!}$ we choose the rankings $\rhd$ that correspond to the extreme points that maximize $\sum_{a_i,a_j \in X} M^{\rhd}_{i,j} \cdot \hat{s}(v, (a_i, a_j))$ subject to $M \in \mathcal{M}$. It is a well-known fact in linear optimization that, if an extreme point is not optimal, then there is a neighboring extreme point with a higher objective value. This insight implies the previous set equality, so we can now infer that $f(R, X) = \hat{g}(v(R), X) \subseteq \{\rhd \in \mathcal{R}(X) : \forall \rhd' \in \mathcal{R}(X) : \hat{s}(v(R), \rhd) \geq \hat{s}(v(R), \rhd')\}$ for all profiles $R \in \mathcal{R}^*$ and all feasible sets $X \in \mathcal{F}(A)$. Or, put differently, f always chooses a subset of the bivariate scoring rule defined by s. As the last step, we use continuity to show that this must be an equality, so f is the bivariate scoring rule induced by s. □

Remark 1 (Independence of the Axioms). All axioms are necessary for Theorem 1. If we only drop local agenda consistency, then every rule that maximizes the score according to a scoring function s that assigns points to every pair of input and output rankings satisfies all axioms, and this class is a strict superset of our bivariate scoring rules. If we only drop reinforcement, we can return the union of the output of two different bivariate scoring rules. Or, to be more precise, one can check that the rule $f(R, X) = f_{\mathit{Borda}}(R, X) \cup f_{\mathit{Kemeny}}(R, X)$ only fails reinforcement. When dropping continuity from our list of axioms, we can use composite bivariate scoring rules which first compute a bivariate scoring rule and then refine it by choosing the rankings with the maximal score according to another bivariate scoring rule. An example of this is the SPF f given by $f(R, X) = \{\rhd \in f_{\mathit{Borda}}(R, X) : \forall \rhd' \in f_{\mathit{Borda}}(R, X) : \hat{s}_{\mathit{Kemeny}}(R, \rhd) \geq \hat{s}_{\mathit{Kemeny}}(R, \rhd')\}$. Faithfulness is required to exclude, e.g., the trivial rule that always returns all rankings or bivariate scoring rules that give a negative score to a pair (a, b) if a voter prefers a to b. To only violate anonymity, we can treat voters differently: for instance, we can count the votes of "even" voters $i \in 2\mathbb{N}$ twice, but the votes of the "odd" voters $i \in 2\mathbb{N} + 1$ only once. Finally, for neutrality, we can, e.g., consider a biased version of Kemeny's rule that counts the points for a specific pair of alternatives twice.

Remark 2 (Faithfulness). The example in Remark 1 demonstrating that faithfulness is required for Theorem 1 generalizes the class of bivariate scoring rules by allowing for bivariate scoring functions s with $s(\ell, k) < 0$ if $\ell < k$ or that $s(\ell, k) = 0$ for all $\ell, k \in \{1, \dots, m\}$. While these rules are not particularly appealing, one may wonder whether it is possible to characterize a more general class of bivariate scoring rules when simply dropping faithfulness. This is not the case because, without faithfulness, local agenda consistency becomes trivial to satisfy: we can simply set $f(R, X) = \mathcal{R}(X)$ for all profiles $R \in \mathcal{R}^*$ and all feasible sets of alternatives $X \in \mathcal{F}(A)$ with $X \neq A$ and define $f(R, A)$ to be any SPF that satisfies anonymity, neutrality, consistency, and, continuity.

Remark 3 (Local Agenda Consistency). It is noteworthy that our proof of Theorem 1 does not use the full power of local agenda consistency. Instead, it suffices to only use local agenda consistency when the smaller feasible set consists of precisely two alternatives. Analogous claims will also hold for Corollaries 1 and 2 and this observation is, e.g., also known for Arrow's impossibility theorem (Schwartz, 1986).

Remark 4 (SSB Utility Functions). Positional scoring rules are closely connected to vNM utility functions: in a positional scoring rule, we assume that each voter has the same canonical vNM utility function and we then order the alternatives according to their social welfare. A similar interpretation is possible for bivariate scoring rule by considering the class of skew-symmetric bilinear (SSB) utility functions (Fishburn, 1984), which assign a value to each pair of alternatives (and thus discard the transitivity enforced by vNM utilities): the decision maker chooses a canonical SSB utility function and the winning rankings are those with maximal social welfare (where the welfare of a ranking is simply the sum of the utilities of each pair of alternatives in the ranking). We note that this interpretation of SSB utility functions is new as SSB utilities have not been studied in the context of ranking aggregation.

3.2. Positional Scoring Rules

Next, we turn to our characterization of positional scoring rules, which we derive as a corollary of Theorem 1 by strengthening local agenda consistency to agenda consistency. We note that Corollary 1 is a variant of the characterization of positional scoring rules by Smith (1973), so our main result intuitively generalizes this prominent result.

Corollary 1. *An SPF is a positional scoring rule if and only if it satisfies anonymity, neutrality, continuity, faithfulness, reinforcement, and agenda consistency.*

Proof. First, we note that all positional scoring rules satisfy anonymity, neutrality, continuity, faithfulness, and reinforcement because they are bivariate scoring rules. Furthermore, these SPFs are agenda consistent because they compute the scores of each alternative with respect to the full profile and order the feasible alternatives in decreasing order of their scores. Hence, if we delete some alternatives from the feasible set, we only need to delete these alternatives from the old rankings to derive the winning rankings for the smaller feasible set.

For the other direction, let f denote an SPF satisfying all given axioms. First, since agenda consistency implies local agenda consistency, we can infer that f is a bivariate scoring rule by Theorem 1. Hence, let s denote the bivariate scoring function of f. We will next show that $s(i, k) = s(i, j) + s(j, k)$ for all $i, j, k \in \{1, \dots, m\}$ with $i < j < k$. Assume for contradiction that this is not the case for some $i, j, k \in \{1, \dots, m\}$ and consider the following profile R for three voters: voter 1 places a at position i, b at position j, and c at position k, voter 2 places b at position i, c at position j, and a at position k, and voter 3 places c at position i, a at position j, and b at position k. All remaining alternatives can be arranged arbitrarily. Next, consider the feasible set $\{a, b\}$: it holds that $\hat{s}(R,(a,b)) = s(i,j)+s(j,k)+s(k,i) \neq 0$ and $\hat{s}(R,(b,a)) = s(j,i)+s(k,j)+s(i,k) = -\hat{s}(R,(a,b))$ because of our assumption that $s(i,k) \neq s(i,j) + s(j,k)$ and the fact that $s(x,y) = -s(y,x)$. This means that $|f(R,\{a,b\})| = 1$ and we suppose without loss of generality that $s(i,j) + s(j,k) + s(k,i) > 0$, so $f(R,\{a,b\}) = \{(a,b)\}$. Based on symmetric calculations, we also get that $f(R,\{b,c\}) = \{(b,c)\}$ and $f(R,\{a,c\}) = \{(c,a)\}$. Finally, agenda consistency implies now for every $\rhd \in f(R,\{a,b,c\})$ that $a \rhd b$, $b \rhd c$, and $c \rhd a$. However, this violates the transitivity of rankings, so we conclude that $s(i,k) = s(i,j) + s(j,k)$ for all $i,j,k \in \{1,\dots,m\}$ with $i < j < k$.

We now define a positional scoring function s' based on s by $s'(i) = s(i,m)$ for all $i \in \{1,\dots,m-1\}$ and $s'(m) = 0$. By our previous insights, it holds that $s(i,j) = s(i,m) - s(j,m)$ for all $i,j \in \{1,\dots,m\}$ with $i < j$. Moreover, if $j < i$, then $s(i,j) = -s(j,i) = -(s(j,m) - s(i,m)) = s(i,m) - s(j,m)$, so the equality holds for all $i,j \in \{1,\dots,m\}$. This means also that s' is non-increasing: if $s'(i) < s'(i+1)$ for some i, then $s(i,i+1) = s(i,m) - s(i+1,m) = s'(i) - s'(i+1) < 0$, which violates the definition of bivariate scoring functions. Finally, since there are $i,j \in \{1,\dots,m\}$ with $s(i,j) > 0$, s' is non-constant. Thus, s' is indeed a positional scoring function.

Next, we note that $\hat{s}(R,(x,y)) = \sum_{i \in N_R} s(r(\succ_i, x), m) - s(r(\succ_i, y), m) = \hat{s}'(R,x) - \hat{s}'(R,y)$ for all profiles R and alternatives $x,y \in A$. This implies for all agendas $X \in \mathcal{F}(A)$ and rankings $\rhd, \rhd' \in \mathcal{R}(X)$ that

$$\begin{aligned}
\hat{s}(R,\rhd) - \hat{s}(R,\rhd') &= \sum_{(x,y)\in\rhd} \hat{s}(R,(x,y)) - \sum_{(x,y)\in\rhd'} \hat{s}(R,(x,y)) \\
&= 2 \sum_{(x,y)\in\rhd\setminus\rhd'} \hat{s}(R,(x,y)) \\
&= 2 \sum_{(x,y)\in\rhd\setminus\rhd'} \hat{s}'(R,x) - \hat{s}'(R,y).
\end{aligned}$$

Finally, let f' denote the positional scoring rule defined by s'; we will show that $f(R,X) = f'(R,X)$ for all profiles $R \in \mathcal{R}^*$ and all feasible sets $X \in \mathcal{F}(A)$. To this end, consider two rankings $\rhd \in f(R,X)$ and $\rhd' \in \mathcal{R}(X) \setminus f(R,X)$ for some $R \in \mathcal{R}^*$ and $X \in \mathcal{F}(A)$. By the definition of f, we have that $\hat{s}(R,\rhd) > \hat{s}(R,\rhd')$, which means that $\sum_{(x,y)\in\rhd\setminus\rhd'} \hat{s}'(R,x) - \hat{s}'(R,y) > 0$. Hence, there is a pair of alternatives $x,y \in X$ such that $\hat{s}'(R,x) > \hat{s}'(R,y)$ and $y \rhd' x$. The definition of positional scoring rules then requires that $\rhd' \notin f'(R,X)$, so $f'(R,X) \subseteq f(R,X)$. For the other direction, consider a ranking $\rhd \in \mathcal{R}(X) \setminus f'(R,X)$. This means that there are alternatives $x,y \in X$ such that $x \rhd y$ but $\hat{s}'(R,x) < \hat{s}'(R,y)$. By iterating through $\rhd$, we can also find two consecutive alternatives $x',y' \in X$ in $\rhd$ that satisfy $x' \rhd y'$ but $\hat{s}'(R,x') < \hat{s}'(R,y')$. Now, consider the ranking $\rhd'$ derived from $\rhd$ by swapping x' and y'. It holds that $\hat{s}(R,\rhd') - \hat{s}(R,\rhd) = 2\hat{s}'(R,y') - 2\hat{s}'(R,x') > 0$, which shows that $\rhd \notin f(R,X)$. Hence, $f(R,X) \subseteq f'(R,X)$ and f is the positional scoring rule induced by s'. $\square$

Remark 5 (Smith's Characterization of Positional Scoring Rules). Corollary 1 is closely related to the characterization of positional scoring rules due to Smith (1973).[5] In more detail, Smith (1973) characterizes positional scoring rules in the context of social welfare functions (SWFs), which return a single weak order over the alternatives rather than a set of strict orders, and shows that an SWF is a positional scoring rule if and only if it satisfies anonymity, neutrality, continuity, and separability. The last axiom, separability, is a variant of reinforcement for SWFs which requires that if $f(R)|_{\{a,b\}} \cap f(R')|_{\{a,b\}} \neq \emptyset$, then $f(R+R')|_{\{a,b\}} = f(R)|_{\{a,b\}} \cap f(R')|_{\{a,b\}}$ for all disjoint profiles $R, R' \in \mathcal{R}^*$ and all alternatives $a, b \in A$ (here, $f(R)$, $f(R')$, and $f(R+R')$ denote the weak rankings on A returned by the SWF f for R, R', and $R+R'$, respectively).

While our result is logically independent to the one by Smith since we work in a different setting, reinforcement and agenda consistency transfer the idea of separability to SPFs. To make this more formal, we define the relation $\succsim_{f(R,X)}$ for all $a, b \in X$ by $a \succsim_{f(R,X)} b$ if and only if there is a ranking $\rhd \in f(R,X)$ with $a \rhd b$. Then, reinforcement and agenda consistency together require for all disjoint profiles $R, R' \in \mathcal{R}^*$, feasible sets $X \in \mathcal{F}(A)$, and alternatives $a, b \in X$ that if $\succsim_{f(R,X)}|_{\{a,b\}} \cap \succsim_{f(R',X)}|_{\{a,b\}} \neq \emptyset$, then $\succsim_{f(R+R',X)}|_{\{a,b\}} = \succsim_{f(R,X)}|_{\{a,b\}} \cap \succsim_{f(R',X)}|_{\{a,b\}}$. The reason for this is that $\succsim_{f(\bar{R},X)}|_{\{a,b\}} = f(\bar{R}, \{a,b\})$ for every profile $\bar{R}$ because of agenda consistency. Hence, if $\succsim_{f(R,X)}|_{\{a,b\}} \cap \succsim_{f(R',X)}|_{\{a,b\}} \neq \emptyset$, then $f(R, \{a,b\}) \cap f(R', \{a,b\}) \neq \emptyset$ and reinforcement shows that $f(R+R', \{a,b\}) = f(R, \{a,b\}) \cap f(R', \{a,b\})$. Applying again agenda consistency then shows that $\succsim_{f(R+R',X)}|_{\{a,b\}} = f(R+R', \{a,b\}) = \succsim_{f(R,X)}|_{\{a,b\}} \cap \succsim_{f(R',X)}|_{\{a,b\}}$, which proves our claim.

3.3. Kemeny's rule

As our last result, we discuss our characterization of Kemeny's rule for which we additionally use independence of infeasible alternatives. Just as Corollary 1, this result follows as a simple corollary from Theorem 1. Since this corollary can be seen as a variant of the characterization of Kemeny's rule by Young (1988), Theorem 1 indeed combines the prominent characterizations of Smith (1973) and Young (1988).

Corollary 2. *An SPF is Kemeny's rule if and only if it satisfies anonymity, neutrality, continuity, faithfulness, reinforcement, local agenda consistency, and independence of infeasible alternatives.*

Proof. Since Kemeny's rule is a bivariate scoring rule, it satisfies all axioms but independence of infeasible alternatives due to Theorem 1. To show that Kemeny's rule also satisfies independence of infeasible alternatives, we consider two profiles R and R' and a feasible set $X \in \mathcal{F}(A)$ such that $N_R = N_{R'}$ and $\succ_i|_X = \succ_i'|_X$ for all $i \in N_R$. This means for all alternatives $x, y \in X$ and all voters $i \in N_R$ that $x \succ_i y$ if and only if $x \succ_i' y$. Furthermore, by the definition of Kemeny's rule, every voter gives 1 point to a ranking

[5] It should be noted that the result by Smith (1973) considers positional scoring rules defined by arbitrary scoring functions s, i.e., this author drops the assumption that s is non-increasing and non-constant. When using agenda consistency instead of local agenda consistency and faithfulness in the proof of Theorem 1, we can also infer this stronger result.

$\rhd \in \mathcal{R}(X)$ for every pair of alternatives $x, y \in X$ with $x \succ_i y$ and $x \rhd y$ and -1 points for every pair of alternatives $x, y \in X$ with $x \succ_i y$ and $y \rhd x$. Combining these two insights implies that $\hat{s}_{Kemeny}(R, \rhd) = \hat{s}_{Kemeny}(R', \rhd)$ for every ranking $\rhd \in \mathcal{R}(X)$ and it thus holds that $f_{Kemeny}(R, X) = f_{Kemeny}(R', X)$.

For the other direction, let f denote an SPF that satisfies all given axioms. By Theorem 1, f is a bivariate scoring rule that satisfies independence of infeasible alternatives and we let s denote its bivariate scoring function. By the definition of bivariate scoring functions, there are indices $i, j \in \{1, \dots, m\}$ with $i < j$ such that $s(i, j) > 0$. Our goal is now to show that $s(i', j') = s(i, j)$ for all indices i', j' with $i' < j'$. To this end, we consider two profiles R and R', which are both defined by two voters. In more detail, in R, a is the i-th best alternative of voter 1 and the j-th best alternative of voter 2, and b is the j-th best alternative of voter 1 and the i-th best alternative of voter 2. All other alternatives can be ranked arbitrarily. Since $s(i, j) = -s(j, i)$, it is easy to check that $f(R, \{a, b\}) = \{(a, b), (b, a)\}$. In the second profile R', voter 2 has the same preference as in R, but voter 1 places a now at position i' and b at position j'. We note that $\succ_k|_{\{a,b\}} = \succ'_k|_{\{a,b\}}$ for $k \in \{1, 2\}$, so independence of infeasible alternatives entails that $f(R', \{a, b\}) = f(R, \{a, b\}) = \{(a, b), (b, a)\}$. This means that $\hat{s}(R', (a, b)) = \hat{s}(R', (b, a))$. Moreover, it holds that $\hat{s}(R', (a, b)) = s(i', j') + s(j, i) = -(s(j', i') + s(i, j)) = -\hat{s}(R', (b, a))$. We therefore infer that $s(i', j') + s(j, i) = s(i', j') - s(i, j) = 0$. Hence, $s(i, j) = s(i', j')$ for all $i', j' \in \{1, \dots, m\}$ with $i' < j'$, which means that f is Kemeny's rule as $s(x, y) = -s(y, x)$ for all $x, y \in \{1, \dots, m\}$ and f is invariant under scaling s. $\square$

Remark 6 (Young's Characterization of Kemeny's Rule). Corollary 2 can be seen as a variant of a prominent characterization in the literature: Young (1988) has shown that Kemeny's rule is the only SPF satisfying anonymity, neutrality, reinforcement, faithfulness, and pairwise consistency (see also Young, 1994, Theorem 6). The last condition in this list, pairwise consistency, requires for all profiles $R \in \mathcal{R}^*$, feasible sets of alternatives $X \in \mathcal{F}(A)$, rankings $\rhd \in f(R, X)$, and alternatives $a, b \in X$ that are consecutive in $\rhd$ that *(i)* if $a \rhd b$, then $(a, b) \in f(R, \{a, b\})$, *(ii)* if $a \rhd b$ and $(b, a) \in f(R, \{a, b\})$, then $(\rhd \setminus \{(a, b)\}) \cup \{(b, a)\} \in f(R, X)$, and *(iii)* $f(R, \{a, b\}) = f(R', \{a, b\})$ for all $R' \in \mathcal{R}^*$ with $N_{R'} = N_R$ and $\succ'_i|_{\{a,b\}} = \succ_i|_{\{a,b\}}$ for all $i \in N_R$. The first condition corresponds to local agenda consistency for agendas of size 2, and the third condition to independence of infeasible alternatives for agendas of size 2. By contrast, none of our axioms relates to the second condition. Since our proofs also work with the restricted notions of local agenda consistency and independence of infeasible alternatives, Corollary 2 gives another variant of this characterization. In particular, our result replaces the second condition of pairwise consistency with continuity.

Remark 7 (Condorcet-consistency). There is another characterization of Kemeny's rule by Young and Levenglick (1978) that relies on a variant of Condorcet-consistency. This axiom postulates for all profiles $R \in \mathcal{R}^*$, feasible sets $X \in \mathcal{F}(A)$, rankings $\rhd \in f(R, X)$, and alternatives $a, b \in X$ that are consecutive in $\rhd$ that *(i)* if $a \rhd b$, then a (weak) majority of voters prefer a to b, and *(ii)* if as many voters prefer a to b

than vice versa and $a \rhd b$, then $(\rhd \setminus \{(a,b)\}) \cup \{(b,a)\} \in f(R,X)$. This definition of Condorcet-consistency differs vastly from classical definitions of Condorcet-consistency and consequently, the result of Young and Levenglick (1978) is often erroneously stated with a weaker variant of Condorcet-consistency.[6] Our main result allows for another variant of this characterization. Firstly, it is easy to see that Kemeny's rule is the only bivariate scoring rule that ranks the Condorcet winner first whenever such an alternative exists. Even more, Corollary 2 also holds when replacing local agenda consistency and independence of infeasible alternatives with another variant of Condorcet-consistency, which requires that if a is ranked directly over b in a winning ranking, then a (weak) majority of the voters must prefer a to b. This axiom implies local agenda consistency and independence of infeasible alternatives for feasible sets of size 2, so this statement follows directly from our proof.

Remark 8 (Arrow's Impossibility Theorem). An immediate consequence of Corollaries 1 and 2 is that no bivariate scoring rule satisfies both agenda consistency and independence of infeasible alternatives. Indeed, when simultaneously demanding both axioms, one can show that Arrow's impossibility theorem applies: only dictatorships satisfy Pareto-optimality, agenda consistency, and independence of infeasible alternatives. This follows analogous to the proofs of Arrow's impossibility in, e.g., Campbell and Kelly (2002). Hence, Corollaries 1 and 2 can be seen as attractive escape routes to this impossibility.

4. Conclusion

In this paper, we study social preference functions (SPFs) which, given the voters' strict preferences over some set of alternatives, compute a non-empty set of winning rankings over a feasible subset of the alternatives. Two of the most prominent classes of SPFs are positional scoring rules and Kemeny's rule, both of which have repeatedly been characterized (e.g., Smith, 1973; Young, 1974b; Young and Levenglick, 1978; Young, 1988). In this paper, we unify these two independent streams of research by characterizing the class of bivariate scoring rules, which contains both Kemeny's rule and all positional scoring rules. Roughly, bivariate scoring rules can be seen as variants of Kemeny's rule that weight comparisons between alternatives differently depending on their positions in the voters' preference relations. We characterize this class of rules by mainly relying on two axioms called reinforcement and local agenda consistency, which formalize consistency notions with respect to variable electorates and variable agendas. Based on this result, we also infer characterizations of the class of positional scoring rules and of Kemeny's rule as corollaries. These characterizations can be seen as variants of the characterizations by Smith (1973) and Young (1988), which demonstrates that our main result indeed combines these two classical theorems. Or, put differently, our main result tries to unify the axiomatic research on positional scoring rules and Kemeny's rule.

[6] Maybe the most prominent occurrence of such a mistake is in the chapter by Zwicker (2016), where Condorcet-consistency only requires that the first-ranked alternative in a winning ranking must be the Condorcet winner if one exists.

Moreover, we note that our characterizations of positional scoring rules and Kemeny's rule also highlight the differences between these SPFs: while the latter SPF satisfies independence of infeasible alternatives, the former ones satisfy a stronger notion of agenda consistency. Since these axioms are jointly incompatible under mild additional conditions due to Arrow's impossibility theorem (Arrow, 1951), these results thus draw a sharp boundary between Kemeny's rule and positional scoring rules.

Finally, we note that our paper can also be seen as progress on the open problem of better understanding SPFs from an axiomatic perspective. Indeed, while there are very general classes of SPFs (see, e.g., Elkind and Slinko, 2016), these large classes are typically not well-understood. By contrast, by focusing on a smaller class of SPFs, we derive a full characterization of an important subset of SPFs.

Acknowledgements

I thank Chris Dong, Dominik Peters, Felix Brandt, and two anonymous reviewers for their helpful suggestions and remarks. This work was supported by the Deutsche Forschungsgemeinschaft under grant BR 2312/12-1.

References

P. Anand. Rationality and intransitive preference: Foundations for the modern view. In P. Anand, P. K. Pattanaik, and C. Puppe, editors, *The Handbook of Rational and Social Choice*, chapter 6. Oxford University Press, 2009.

K. J. Arrow. *Social Choice and Individual Values*. New Haven: Cowles Foundation, 1st edition, 1951. 2nd edition 1963.

K. J. Arrow, A. K. Sen, and K. Suzumura, editors. *Handbook of Social Choice and Welfare*, volume 1. North-Holland, 2002.

M. Bar-Hillel and A. Margalit. How vicious are cycles of intransitive choice? *Theory and Decision*, 24(2):119–145, 1988.

J. P. Barthelemy and B. Monjardet. The median procedure in cluster analysis and social choice theory. *Mathematical Social Sciences*, 1:235–267, 1981.

J. Bartholdi, III, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6(3):227–241, 1989.

N. Boehmer, R. Bredereck, and D. Peters. Rank aggregation using scoring rules. Technical report, https://arxiv.org/pdf/2209.08856, 2023.

G. Bordes. Consistency, rationality and collective choice. *Review of Economic Studies*, 43(3):451–457, 1976.

W. Bossert and Y. Sprumont. Strategy-proof preference aggregation: Possibilities and characterizations. *Games and Economic Behavior*, 85:109–126, 2014.

F. Brandl and D. Peters. Approval voting under dichotomous preferences: A catalogue of characterizations. *Journal of Economic Theory*, 205, 2022.

F. Brandl, F. Brandt, and H. G. Seedig. Consistent probabilistic social choice. *Econometrica*, 84(5):1839–1880, 2016.

F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

D. E. Campbell and J. S. Kelly. Impossibility theorems in the Arrovian framework. In K. J. Arrow, A. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 1, chapter 1. Elsevier, 2002.

B. Can. Weighted distances between preferences. *Journal of Mathematical Economics*, 51:109–115, 2014.

B. Can and T. Storcken. Update monotone preference rules. *Mathematical Social Sciences*, 65(2):136–149, 2013.

D. P. Cervone, W. V. Gehrlein, and W. S. Zwicker. Which scoring rule maximizes Condorcet efficiency under IAC? *Theory and Decision*, 58:145–185, 2005.

P. Y. Chebotarev and E. Shamis. Characterizations of scoring methods for preference aggregation. *Annals of Operations Research*, 80:299–332, 1998.

S. Ching. A simple characterization of plurality rule. *Journal of Economic Theory*, 71 (1):298–302, 1996.

V. Conitzer, M. Rognlie, and L. Xia. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 109–115, 2009.

W. D. Cook and M. Kress. Ordinal ranking and preference strength. *Mathematical Social Sciences*, 11:295–306, 1986.

C. Dong and P. Lederer. Refined characterizations of approval-based committee scoring rules. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, 2024. Forthcoming.

C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank aggregation methods for the web. In *Proceedings of the 10th International Conference on the World Wide Web (WWW)*, pages 613–622, 2001.

E. Elkind and A. Slinko. Rationalizations of voting rules. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 8. Cambridge University Press, 2016.

P. Favardin and D. Lepelley. Some further results on the manipulability of social choice rules. *Social Choice and Welfare*, 26(3):485–509, 2006.

P. C. Fishburn. The irrationality of transitity in social choice. *Behavioral Science*, 15: 119–123, 1970.

P. C. Fishburn. Axioms for approval voting: Direct proof. *Journal of Economic Theory*, 19(1):180–185, 1978.

P. C. Fishburn. SSB utility theory: An economic perspective. *Mathematical Social Sciences*, 8(1):63–94, 1984.

P. C. Fishburn and W. V. Gehrlein. Borda's rule, positional voting, and Condorcet's simple majority principle. *Public Choice*, 28(1):79–88, 1976.

W. Gaertner. Kenneth Arrow's impossibility theorem stretching to other fields. *Public Choice*, 179(1–2):125–131, 2019.

W. V. Gehrlein. Condorcet efficiency and constant scoring rules. *Mathematical Social Sciences*, 2(2):123–130, 1982.

P. C. Gilmore and A. J. Hoffmann. A characterization of comparability graphs and of interval graphs. *Canadian Journal of Mathematics*, 16:539–548, 1964.

E. Hemaspaandra, H. Spakowski, and J. Vogel. The complexity of Kemeny elections. *Theoretical Computer Science*, 349(3):382–391, 2005.

J. G. Kemeny. Mathematics without numbers. *Daedalus*, 88:577–591, 1959.

C. Klamler. The Dodgson ranking and its relation to Kemeny's method and Slater's rule. *Social Choice and Welfare*, 23:91–102, 2004.

R. Kolde, S. Laur, P. Adler, and J. Vilo. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics*, 2012.

R. Kumar and S. Vassilvitskii. Generalized distances between rankings. In *Proceedings of the 19th International Conference on the World Wide Web (WWW)*, pages 571–580, 2010.

M. Lackner and P. Skowron. Consistent approval-based multi-winner rules. *Journal of Economic Theory*, 192:105173, 2021.

M. Lackner and P. Skowron. Multi-winner voting with approval preferences. https://arxiv.org/abs/2007.01795, 2022.

G. Laffond, J. Lainé, and J.-F. Laslier. Composition-consistent tournament solutions and social choice functions. *Social Choice and Welfare*, 13(1):75–93, 1996.

J.-F. Laslier. *Tournament Solutions and Majority Voting*. Springer-Verlag, 1997.

S. Lin. Rank aggregation methods. *WIREs computational statistics*, 2010.

K. May. Intransitivity, utility, and the aggregation of preference patters. *Econometrica*, 22(1):1–13, 1954.

R. B. Myerson. Axiomatic derivation of scoring rules without the ordering assumption. *Social Choice and Welfare*, 12(1):59–74, 1995.

S. Nitzan and A. Rubinstein. A further characterization of Borda ranking method. *Public Choice*, 36(1):153–158, 1981.

M. Pivato. Variable-population voting rules. *Journal of Mathematical Economics*, 49 (3):210–221, 2013.

M. Pivato. Asymptotic utilitarianism in scoring rules. *Social Choice and Welfare*, 47(2): 431–458, 2016.

A. Plaia, S. Buscemi, and M. Sciandra. A new position weight correlation coefficient for consensus ranking process without ties. *Stat*, 8:e236, 2019.

R. C. Prati. Combining feature ranking algorithms through rank aggregation. In *International Conference on Neural Networks*, pages 1–8, 2012.

G. Pritchard and M. C. Wilson. Exact results on the manipulability of positional scoring rules. *Social Choice and Welfare*, 29:487–513, 2007.

T. C. Ratliff. A comparison of Dodgson's method and Kemeny's rule. *Social Choice and Welfare*, 18:79–89, 2001.

M. E. Renda and U. Straccia. Web metasearch: rank vs score based rank aggregation methods. In *ACM Symposium on Applied Computing*, pages 841–846, 2003.

J. Richelson. A comparative analysis of social choice functions, III. *Behavioral Science*, 23(3):169–176, 1978.

R. T. Rockafeller. *Convex Analysis*. Princeton University Press, 1970.

D. G. Saari and V. Merlin. A geometric examination of kemeny's rule. *Social Choice and Welfare*, 17(3):403–438, 2000.

C. Sarkar, S. Cooley, and J. Srivastava. Robust feature selection technique using rank aggregation. *Applied Artificial Intelligence*, 28(3):243–257, 2014.

T. Schwartz. *The Logic of Collective Choice*. Columbia University Press, 1986.

A. Sen. Choice functions and revealed preference. *Review of Economic Studies*, 38(3): 307–317, 1971.

A. K. Sen. Social choice theory: A re-examination. *Econometrica*, 45(1):53–89, 1977.

A. K. Sen. *Collective Choice and Social Welfare*. Penguin, expanded edition, 2017.

P. Skowron and E. Elkind. Social choice under metric preferences: scoring rules and STV. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 706–712, 2017.

P. Skowron, P. Faliszewski, and A. Slinko. Axiomatic characterization of committee scoring rules. *Journal of Economic Theory*, 180:244–273, 2019.

P. Slater. Inconsistencies in a schedule of paired comparisons. *Biometrika*, 48(3–4): 303–312, 1961.

J. H. Smith. Aggregation of preferences with variable electorate. *Econometrica*, 41(6): 1027–1041, 1973.

T. N. Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206, 1987.

H. P. Young. An axiomatization of Borda's rule. *Journal of Economic Theory*, 9(1): 43–52, 1974a.

H. P. Young. A note on preference aggregation. *Econometrica*, 42(6):1129–1131, 1974b.

H. P. Young. Social choice scoring functions. *SIAM Journal on Applied Mathematics*, 28(4):824–838, 1975.

H. P. Young. On permutations and permutation polytopes. In M. L. Balinski and A. J. Hoffman, editors, *Polyhedral Combinatorics: Dedicated to the memory of D.R. Fulkerson*, pages 128–140. Springer Berlin Heidelberg, 1978.

H. P. Young. Condorcet's theory of voting. *The American Political Science Review*, 82 (4):1231–1244, 1988.

H. P. Young. *Equity: in Theory and Practice*. Princeton University Press, 1994.

H. P. Young. Optimal voting rules. *Journal of Economic Perspectives*, 9(1):51–64, 1995.

H. P. Young and A. Levenglick. A consistent extension of Condorcet's election principle. *SIAM Journal on Applied Mathematics*, 35(2):285–300, 1978.

W. S. Zwicker. Consistency without neutrality in voting rules: When is a vote an average? *Mathematical and Computer Modelling*, 48(9):1357 – 1373, 2008.

W. S. Zwicker. Introduction to the theory of voting. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 2. Cambridge University Press, 2016.

A. Appendix: Proof of Theorem 1

In this appendix, we provide a complete proof of Theorem 1: an SPF is a bivariate scoring rule if and only if it satisfies anonymity, neutrality, continuity, faithfulness, reinforcement, and local agenda consistency. Since we have shown in the main body that every bivariate scoring rule satisfies the given axioms, we focus here on the converse and thus assume throughout the appendix that f is an SPF that satisfies all six axioms. Our goal is to find the underlying bivariate scoring function of f, and we will follow the proof sketch in the main body for this. For a better readability, we organize the proof in several lemmas, which are grouped by subsections: in Appendix A.1 we show that f is non-imposing, in Appendix A.2 we apply the separating hyperplane theorem for convex sets to infer the vectors $u^{\rhd,\rhd'}$, in Appendix A.3 we study feasible sets of size 2, and in Appendix A.4, we extend our reasoning to larger feasible sets and prove Theorem 1.

A.1. Non-imposition

Our first goal is to show that the SPF f is non-imposing, i.e., that there is for every feasible set $X \in \mathcal{F}(A)$ and every ranking $\rhd \in \mathcal{R}(X)$ a profile R such that $f(R, X) = \{\rhd\}$. To this end, we first construct an auxiliary profile R^a, in which all rankings $\rhd \in \mathcal{R}(X)$ are chosen that top-rank a given alternative $a \in X$.

Lemma 1. *For all feasible sets $X \in \mathcal{F}(A)$ and alternatives $a \in X$, there is a profile R^a such that $f(R^a, X) = \{\rhd \in \mathcal{R}(X) \colon \forall x \in X \setminus \{a\} \colon a \rhd x\}$.*

Proof. Fix an arbitrary feasible set $X \in \mathcal{F}(A)$ and an alternative $a \in X$. If $|X| = 1$, the lemma is trivial and we thus suppose that $|X| \geq 2$. For constructing the profile R^a, we will first derive a profile $R^{a,b}$ such that $f(R^{a,b}, \{a, b\}) = \{(a, b)\}$ for some alternative $b \in X \setminus \{a\}$, $(a, x) \in f(R^{a,b}, \{a, x\})$ for all $x \in X \setminus \{a, b\}$, and $f(R^{a,b}, \{x, y\}) = \{(x, y), (y, x)\}$ for all $x, y \in X \setminus \{a, b\}$. To this end, we recall that there is a preference relation $\succ^* \in \mathcal{R}(A)$ such that $a \succ^* b$ and $f(\succ^*, \{a, b\}) = \{(a, b)\}$ by faithfulness.

The profile $R^{a,b}$ contains $\frac{m!}{2}$ voters such that every preference relation $\succ \in \mathcal{R}(A)$ with $a \succ b$ is reported by one voter. An example of $R^{a,b}$ for $m = 4$ alternatives is shown in Figure 2. By faithfulness, it holds for every preference relation $\succ$ in $R^{a,b}$ that $(a, b) \in f(\succ, \{a, b\})$. Since $\succ^*$ also appears in $R^{a,b}$, we can infer from reinforcement that $f(R^{a,b}, \{a, b\}) = \{(a, b)\}$. Furthermore, we note that in $R^{a,b}$, all alternatives $x, y \in X \setminus \{a, b\}$ are completely symmetric. In more detail, it holds for every permutation $\tau : A \to A$ with $\tau(a) = a$ and $\tau(b) = b$ that the profile $\tau(R)$ (that is derived by permuting the alternatives in R according to τ) is the same as R up to renaming the voters. As a consequence, anonymity and neutrality require that $f(R^{a,b}, \{x, y\}) = \{(x, y), (y, x)\}$ for all $x, y \in X \setminus \{a, b\}$.

Finally, we will show that $(a, x) \in f(R^{a,b}, \{a, x\})$ for all $x \in X \setminus \{a, b\}$. To this end, we fix an alternative $x \in A \setminus \{a, b\}$ and let $\tau : A \to A$ denote the permutation that only exchanges a and x, i.e., $\tau(a) = x$, $\tau(x) = a$, and $\tau(z) = z$ for all $z \in A \setminus \{a, x\}$. The central insight for this claim is that for every preference relation $\succ$ in $R^{a,b}$ with $x \succ a$, the preference relation $\tau(\succ)$ is also in $R^{a,b}$ because x is preferred to b whenever it is preferred

$R^{a,b}$	1: c, a, b, d	1: c, a, d, b	1: c, d, a, b	1: d, c, a, b
	1: a, c, b, d	1: a, c, d, b	1: a, d, c, b	1: d, a, c, b
	1: a, b, c, d	1: a, b, d, c	1: a, d, b, c	1: d, a, b, c

Figure 2: The profile $R^{a,b}$ constructed in the proof of Lemma 1 for the case that $A = \{a, b, c, d\}$. The k-th row (for $k \in \{1, 2, 3\}$) shows the voters in the set N^k when $x = c$.

to a. We thus partition the voters $i \in N_{R^{a,b}}$ into three sets: $N^1 = \{i \in N_{R^{a,b}} : x \succ_i a \succ_i b\}$ contains the voters that prefer x to a to b, $N^2 = \{i \in N_{R^{a,b}} : a \succ_i x \succ_i b\}$ is the set of voters that prefer a to x to b, and $N^3 = \{i \in N_{R^{a,b}} : a \succ_i b \succ_i x\}$ is the set of voters that prefer a to b to x. Moreover, we define the profiles R^k for $k \in \{1, 2, 3\}$ as the restriction of $R^{a,b}$ to N^k, i.e., R^k is given by $N_{R^k} = N^k$ and $\succ_i^k = \succ_i^{a,b}$ for all $i \in N^k$. First, we note that $R^1 + R^2 + R^3 = R^{a,b}$. Secondly, by our previous observation, $\tau(R^1) = R^2$ and $\tau(R^2) = R^1$ (except for renaming voters), so we can conclude that $f(R^1 + R^2, \{a, x\}) = \{(a, x), (x, a)\}$ due to anonymity and neutrality. Finally, faithfulness and reinforcement imply that $(a, x) \in f(R^3, \{a, x\})$ because every voter prefers a to x in R^3. Hence, we can apply reinforcement again to derive that $(a, x) \in f(R^1 + R^2, \{a, x\}) \cap f(R^3, \{a, x\}) = f(R^{a,b}, \{a, x\})$.

The profile R^a consists now of a copy of $R^{a,b}$ for every alternative $b \in X \setminus \{a\}$, i.e., $R^a = \sum_{b \in X \setminus \{a\}} R^{a,b}$. By our previous analysis, we have that $f(R^{a,b}, \{a, b\}) = \{(a, b)\}$ and $(a, x) \in f(R^{a,b}, \{a, x\})$ for all $b \in X \setminus \{a\}$, $x \in X \setminus \{a, b\}$. Hence, reinforcement shows that $f(R^a, \{a, x\}) = \{(a, x)\}$ for all alternatives $x \in X \setminus \{a\}$. By local agenda consistency, this means that $a \rhd x$ for all rankings $\rhd \in f(R^a, X)$ and alternatives $x \in X \setminus \{a\}$; otherwise, there is a ranking $\rhd' \in f(R^a, X)$ and an alternative $x \in X$ such that $x \rhd' a$ and a, x are consecutive in $\rhd'$. However, local agenda consistency then implies that $(x, a) \in f(R^a, \{a, x\})$, which contradicts our previous analysis. So, $f(R^a, X) \subseteq \{\rhd \in \mathcal{R}(X) : \forall x \in X \setminus \{a\} : a \rhd x\}$. Finally, all alternatives in $X \setminus \{a\}$ are symmetric in R^a. In more detail, for every permutation $\tau : A \to A$ with $\tau(a) = a$ and $\tau(x) = x$ for $x \in A \setminus X$, it holds that $\tau(R^a)$ and R^a differ only in the naming of the voters because, for every $x \in X \setminus \{a\}$, $\tau(R^{a,x})$ is equal to $R^{a,\tau(x)}$ up to renaming the voters. Hence, $\tau(R^a) = \sum_{x \in X \setminus \{a\}} \tau(R^{a,x}) = \sum_{x \in X \setminus \{a\}} R^{a,\tau(x)} = \sum_{x \in X \setminus \{a\}} R^{a,x} = R^a$ for all such permutations τ. Consequently, anonymity and neutrality require for every ranking $\rhd \in f(R^a)$ and every permutation $\tau : A \to A$ with $\tau(x) = x$ for $x \in (A \setminus X) \cup \{a\}$ that $\tau(\rhd) \in f(R^a)$, which proves that $f(R^a, X) = \{\rhd \in \mathcal{R}(X) : \forall x \in X \setminus \{a\} : a \rhd x\}$. $\square$

Based on Lemma 1, we next show that f is non-imposing.

Lemma 2. *The SPF f is non-imposing.*

Proof. Fix an arbitrary feasible set $X = \{a_1, \dots, a_{|X|}\} \in \mathcal{F}(A)$ and a ranking $\rhd = (a_1, a_2, \dots, a_{|X|}) \in \mathcal{R}(X)$. We will inductively derive profiles R^ℓ such that $f(R^\ell, X) = \{\rhd \in \mathcal{R}(X) : \forall x \in X \setminus \{a_1, \dots, a_\ell\} : a_1 \rhd \dots \rhd a_\ell \rhd x\}$ for all $\ell \in \{1, \dots, |X| - 1\}$. Less formally, f should return for R^ℓ all preference relations on X which rank the first

ℓ alternatives like $\rhd$ and the remaining alternatives can be placed in every possible way. Furthermore, all alternatives $x \in X \setminus \{a_1, \dots, a_\ell\}$ will be completely symmetric in the profile R^ℓ, i.e., for every permutation $\tau : A \to A$ with $\tau(y) = y$ for all $y \in A \setminus \{a_{\ell+1}, \dots, a_{|X|}\}$, it holds that R^ℓ is equal to $\tau(R^\ell)$ up to renaming the voters. This induction proves our lemma because $f(R^\ell, X) = \{\rhd\}$ when $\ell = |X| - 1$.

The induction basis $\ell = 1$ follows from Lemma 1 as $f(R^{a_1}, X) = \{\rhd \in \mathcal{R}(X)\colon \forall x \in X \setminus \{a_1\}\colon a_1 \rhd x\}$. Hence, suppose that there is $\ell \in \{1, \dots, |X| - 2\}$ such that $f(R^\ell, X) = \{\rhd \in \mathcal{R}(X)\colon \forall x \in X \setminus \{a_1, \dots, a_\ell\}\colon a_1 \rhd \dots \rhd a_\ell \rhd x\}$. Moreover, we suppose that R^ℓ satisfies the symmetry condition specified before. As the first point, we observe that $f(R^\ell, \{a_{\ell+1}, \dots, a_{|X|}\}) = \mathcal{R}(\{a_{\ell+1}, \dots, a_{|X|}\})$ because of local agenda consistency, i.e., all rankings on $\{a_{\ell+1}, \dots, a_{|X|}\}$ are chosen in $f(R^\ell, \{a_{\ell+1}, \dots, a_{|X|}\})$. Next, we consider the profile $R^{a_{\ell+1}}$ of Lemma 1 for which $f(R^{a_{\ell+1}}, \{a_{\ell+1}, \dots, a_{|X|}\}) = \{\rhd \in \mathcal{R}(\{a_{\ell+1}, \dots, a_{|X|}\})\colon \forall x \in \{a_{\ell+2}, \dots, a_{|X|}\}\colon a_{\ell+1} \rhd x\}$. By reinforcement, we derive for every $\lambda \in \mathbb{N}$ that $f(\lambda R^\ell + R^{a_{\ell+1}}, \{a_{\ell+1}, \dots, a_{|X|}\}) = \{\rhd \in \mathcal{R}(\{a_{\ell+1}, \dots, a_{|X|}\})\colon \forall x \in \{a_{\ell+2}, \dots, a_{|X|}\}\colon a_{\ell+1} \rhd x\}$. Moreover, continuity implies that there is a $\lambda \in \mathbb{N}$ such that $f(\lambda R^\ell + R^{a_{\ell+1}}, X) \subseteq f(R^\ell, X)$. Now, if there is a ranking $\rhd \in f(\lambda R^\ell + R^{a_{\ell+1}}, X)$ such that $a_j \rhd a_{\ell+1}$ for $j > \ell + 1$, local agenda consistency entails that $\rhd|_{\{a_{\ell+1}, \dots, a_{|X|}\}} \in f(\lambda R^\ell + R^{a_{\ell+1}}, \{a_{\ell+1}, \dots, a_{|X|}\})$. However, this contradicts that $f(\lambda R^\ell + R^{a_{\ell+1}}, \{a_{\ell+1}, \dots, a_{|X|}\}) = \{\rhd \in \mathcal{R}(\{a_{\ell+1}, \dots, a_{|X|}\})\colon \forall x \in \{a_{\ell+2}, \dots, a_{|X|}\}\colon a_{\ell+1} \rhd x\}$, so it follows that $f(\lambda R^\ell + R^{a_{\ell+1}}, X) \subseteq \{\rhd \in \mathcal{R}(X)\colon \forall x \in X \setminus \{a_1, \dots, a_{\ell+1}\}\colon a_1 \rhd \dots \rhd a_\ell \rhd a_{\ell+1} \rhd x\}$.

Finally, we note that the alternatives in $\{a_{\ell+2}, \dots, a_{|X|}\}$ are completely symmetric in both R^ℓ and $R^{a_{\ell+1}}$ (i.e., for every permutation $\tau : A \to A$ with $\tau(y) = y$ for all $y \in A \setminus \{a_{\ell+2}, \dots, a_k\}$, the profile R^ℓ and $R^{a_{\ell+1}}$ are equal to $\tau(R^\ell)$ and $\tau(R^{a_{\ell+1}})$ up to renaming the voters): for R^ℓ, this is follows from the induction hypothesis, and for $R^{a_{\ell+1}}$, a careful inspection of the proof of Lemma 1 shows this claim. Therefore, anonymity and neutrality require for every ranking $\rhd \in f(\lambda R^\ell + R^{a_{\ell+1}}, X)$ and every permutation $\tau : A \to A$ with $\tau(y) = y$ for all $y \in A \setminus \{a_{\ell+2}, \dots, a_{|X|}\}$ that $\tau(\rhd) \in f(\lambda R^\ell + R^{a_{\ell+1}}, X)$. Hence, $f(R^{\ell+1}, X) = \{\rhd \in \mathcal{R}(X)\colon \forall x \in X \setminus \{a_1, \dots, a_{\ell+1}\}\colon a_1 \rhd \dots \rhd a_{\ell+1} \rhd x\}$ for the profile $R^{\ell+1} = \lambda R^\ell + R^{a_{\ell+1}}$. Moreover, since R^ℓ and $R^{a_{\ell+1}}$ are symmetric for the alternatives $\{a_{\ell+2}, \dots, a_k\}$, the profile $R^{\ell+1}$ satisfies all our conditions and it therefore proves the induction step as well as the lemma. $\square$

A.2. Separating Hyperplanes

After establishing that f is non-imposing, we will work towards deriving the bivariate scoring function of f. For doing so, we will use the separating hyperplane theorem for convex sets as, e.g., showcased by Young (1975) and Young and Levenglick (1978). For this, we will first change the representation of f from preference profiles to a numerical space. We hence assume that there is a bijection $b : \{1, \dots, m!\} \to \mathcal{R}(A)$ that enumerates all preference relations and denote by $\succ^k$ the preference relation with $b(k) = \succ^k$. Based on this enumeration, we can represent preference profiles as as vectors $v \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$:

the entry v_k states how often the preference relation $\succ^k$ appears in the given profile.[7] To indicate the vector corresponding to a profile R, we will write $v(R)$, so $v(R)_k$ states how many voters report the preference relation $\succ^k$ in R. Now, since f is anonymous, there is a (unique) function $g : \mathbb{N}_0^{m!} \setminus \{\vec{0}\} \times \mathcal{F}(A) \rightarrow \bigcup_{X \in \mathcal{F}(A)} \mathcal{F}(\mathcal{R}(X))$ such that $f(R, X) = g(v(R), X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$. Moreover, it is easy to see that g inherits all desirable properties of f:

- Neutrality: $g(\tau(v), \tau(X)) = \{\tau(\rhd) \colon \rhd \in g(v, X)\}$ for all vectors $v \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$, feasible sets $X \in \mathcal{F}(A)$, and permutations $\tau : A \rightarrow A$. Here, $\tau(v)$ denotes the vector defined by $\tau(v)_\ell = v_k$ for all indices ℓ, k with $\tau(\succ^k) = \succ^\ell$. This ensures that $v(\tau(R)) = \tau(v(R))$.
- Reinforcement: $g(v + v', X) = g(v, X) \cap g(v', X)$ for all vectors $v, v' \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$ and feasible sets $X \in \mathcal{F}(A)$ with $g(v, X) \cap g(v', X) \neq \emptyset$.
- Local agenda consistency: $g(v, X) \supseteq \{\rhd|_X \colon \rhd \in g(v, Y) \cap C(X, Y)\}$ for all $v \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$ and all feasible sets $X, Y \in \mathcal{F}(A)$ with $X \subseteq Y$.

We next extend g to a function $\hat{g}$ on the domain $\mathbb{Q}^{m!} \times \mathcal{F}(A)$ that also satisfies all desirable properties of f (i.e., $\hat{g}$ satisfies these properties for all vectors $v \in \mathbb{Q}^{m!}$ and feasible sets $X \in \mathcal{F}(A)$).

Lemma 3. *There is a neutral, reinforcing, and locally agenda consistent function $\hat{g} : \mathbb{Q}^{m!} \times \mathcal{F}(A) \rightarrow \bigcup_{X \in \mathcal{F}(A)} \mathcal{F}(\mathcal{R}(X))$ such that $\hat{g}(v, X) \subseteq \mathcal{R}(X)$ for all vectors $v \in \mathbb{Q}^{m!}$ and feasible sets $X \in \mathcal{F}(A)$ and $f(R, X) = \hat{g}(v(R), X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$.*

Proof. As discussed before the lemma, there is a function $g : \mathbb{N}_0^{m!} \setminus \{\vec{0}\} \times \mathcal{F}(A) \rightarrow \bigcup_{X \in \mathcal{F}(A)} \mathcal{F}(\mathcal{R}(X))$ that satisfies the requirements of the lemma. We will extend the domain of this function from $\mathbb{N}_0^{m!} \setminus \{\vec{0}\} \times \mathcal{F}(A)$ to $\mathbb{Q}^{m!} \times \mathcal{F}(A)$. To this end, we rely on the profile R^* on $m!$ voters in which every preference relation $\succ \in \mathcal{R}(A)$ is reported by exactly one voter. Moreover, we define $v^* = v(R^*)$ and note that $f(R^*, X) = g(v^*, X) = \mathcal{R}(X)$ for every feasible set $X \in \mathcal{F}(A)$ because of anonymity and neutrality. We will first generalize g to $\mathbb{Z}^{m!}$ and then to $\mathbb{Q}^{m!}$.

Step 1: Extension to $\mathbb{Z}^{m!}$

For this step, we define the function $\bar{g}(v, X) = g(v + \lambda v^*, X)$ for all $v \in \mathbb{Z}^{m!}$ and $X \in \mathcal{F}(A)$, where $\lambda \in \mathbb{N}_0$ is an arbitrary scalar such that $v + \lambda v^* \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$. First, we note that $\bar{g}$ is defined for all $v \in \mathbb{Z}^{m!}$ because for all such vectors, there is $\lambda \in \mathbb{N}_0$ such that $v + \lambda v^* \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$. Next, we show that $\bar{g}$ is well-defined despite the fact that we do not fully specify λ. To see this, consider a feasible set $X \in \mathcal{F}(A)$, a vector $v \in \mathbb{Z}^{m!}$, and two values $\lambda_1, \lambda_2 \in \mathbb{N}_0$ such that $v + \lambda_1 v^* \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$ and $v + \lambda_2 v^* \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$. We will show that $g(v + \lambda_1 v^*, X) = g(v + \lambda_2 v^*, X)$, which entails that $\bar{g}$ is well-defined. If $\lambda_1 = \lambda_2$, this is obvious. Hence, suppose without loss of generality that $\lambda_1 > \lambda_2$, which implies

[7] $\mathbb{N}_0$ is the set $\mathbb{N} \cup \{0\}$ and $\vec{0}$ denotes the vector that contains only 0 entries. Since preference profiles contain always at least one preference relation, no preference profile is mapped to $\vec{0}$.

that $v + \lambda_1 v^* = v + \lambda_2 v^* + (\lambda_1 - \lambda_2)v^*$. Since $g((\lambda_1 - \lambda_2)v^*, X) = \mathcal{R}(X)$, we infer by reinforcement that $g(v+\lambda_1 v^*, X) = g(v+\lambda_2 v^*, X) \cap g((\lambda_1 - \lambda_2)v^*, X) = g(v+\lambda_2 v^*, X)$, which proves our claim. Furthermore, we note that $f(R, X) = g(v(R) + 0v^*, X) = \bar{g}(v(R), X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$.

Next, we show that $\bar{g}$ satisfies all required properties. To this end, consider an arbitrary vector $v_1 \in \mathbb{Z}^{m!}$ and a scalar $\lambda_1 \in \mathbb{N}_0$ such that $v_1 + \lambda_1 v^* \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$. First, for neutrality, we note that $\tau(v_1 + \lambda_1 v^*) = \tau(v_1) + \lambda_1 v^*$ for every permutation $\tau : A \to A$ because $\tau(\lambda_1 v^*) = \lambda_1 v^*$. This implies that $\bar{g}(\tau(v_1), \tau(X)) = g(\tau(v_1) + \lambda_1 v^*, \tau(X)) = g(\tau(v_1 + \lambda_1 v^*), \tau(X)) = \{\tau(\rhd) \colon \rhd \in g(v_1 + \lambda_1 v^*, X)\} = \{\tau(\rhd) \colon \rhd \in \bar{g}(v_1, X)\}$ because of the neutrality of g, so $\bar{g}$ also satisfies this axiom.

As the second axiom, we consider local agenda consistency and consider therefore two feasible sets $X, Y \in \mathcal{F}(A)$ with $X \subseteq Y$. By the local agenda consistency of g, it follows that $\bar{g}(v_1, X) = g(v_1 + \lambda_1 v^*, X) \supseteq \{\rhd|_X \colon \rhd \in g(v_1 + \lambda_1 v^*, Y) \cap C(X, Y)\} = \{\rhd|_X \colon \rhd \in \bar{g}(v_1, Y) \cap C(X, Y)\}$, which demonstrates that $\bar{g}$ satisfies this axiom, too.

Finally, for reinforcement, let $v_2 \in \mathbb{Z}^{m!}$ denote another vector and $\lambda_2 \in \mathbb{N}_0$ such that $v_2 + \lambda_2 v^* \in \mathbb{N}_0^{m!} \setminus \{\vec{0}\}$ and suppose that $\bar{g}(v_1, X) \cap \bar{g}(v_2, X) \neq \emptyset$ for some $X \in \mathcal{F}(A)$. Since g is reinforcing, it holds that $\bar{g}(v_1 + v_2, X) = g(v_1 + v_2 + (\lambda_1 + \lambda_2)v^*, X) = g(v_1 + \lambda_1 v^*, X) \cap g(v_2 + \lambda_2 v^*, X) = \bar{g}(v_1, X) \cap \bar{g}(v_2, X)$. Thus, $\bar{g}$ also satisfies reinforcement.

Step 2: Extension to $\mathbb{Q}^{m!}$

In the second step, we will extend $\bar{g}$ to $\mathbb{Q}^{m!}$. To this end, we define $\hat{g}(v, X) = \bar{g}(\lambda v, X)$ for all vectors $v \in \mathbb{Q}^{m!}$ and feasible sets $X \in \mathcal{F}(A)$, where $\lambda \in \mathbb{N}$ is an arbitrary scalar such that $\lambda v \in \mathbb{Z}^{m!}$. First, we note that $\bar{g}$ is defined for all $v \in \mathbb{Q}^{m!}$ because each such vector can be represented as $\frac{v'}{\lambda}$ for some $v' \in \mathbb{Z}^{m!}$ and $\lambda \in \mathbb{N}$. Next, we will show that $\hat{g}$ is well-defined. For this, we consider a vector $v \in \mathbb{Q}^{m!}$ and two scalars $\lambda_1, \lambda_2 \in \mathbb{N}$ such that $\lambda_1 v \in \mathbb{Z}^{m!}$ and $\lambda_2 v \in \mathbb{Z}^{m!}$. Since $\bar{g}$ is reinforcing, it holds that $\bar{g}(\lambda_1 v, X) = \bar{g}(\lambda_1 \lambda_2 v, X) = \bar{g}(\lambda_2 v, X)$, so $\hat{g}$ is well-defined. Moreover, we observe that $f(R, X) = \bar{g}(1 \cdot v(R), X) = \hat{g}(v(R), X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$.

As the last point of this lemma, we again verify that $\hat{g}$ satisfies all required axioms. To this end, we let $v_1 \in \mathbb{Q}^{m!}$ denote an arbitrary vector and $\lambda_1 \in \mathbb{N}$ denote a scalar such that $\lambda_1 v_1 \in \mathbb{Z}^{m!}$. Now, for showing that $\hat{g}$ is neutral, we note that $\lambda_1 v_1 \in \mathbb{Z}^{m!}$ implies that $\lambda_1 \tau(v_1) \in \mathbb{Z}^{m!}$ for every permutation $\tau : A \to A$. Hence, it holds that $\hat{g}(\tau(v_1), \tau(X)) = \bar{g}(\lambda_1 \tau(v_1), \tau(X)) = \bar{g}(\tau(\lambda_1 v_1), \tau(X)) = \{\tau(\rhd) \colon \bar{g}(\lambda_1 v_1, X)\} = \{\tau(\rhd) \colon \hat{g}(v_1, X)\}$ for every feasible set $X \in \mathcal{F}(A)$ and permutation $\tau : A \to A$, so $\hat{g}$ is neutral.

Next, we show that $\hat{g}$ is locally agenda consistent, for which we consider two feasible sets $X, Y \in \mathcal{F}(A)$ with $X \subseteq Y$. It is easy to see that $\hat{g}(v_1, X) = \bar{g}(\lambda_1 v_1, X) \supseteq \{\rhd|_X \colon \rhd \in \bar{g}(\lambda_1 v_1, Y) \cap C(X, Y)\} = \{\rhd|_X \colon \rhd \in \hat{g}(v_1, Y) \cap C(X, Y)\}$ because $\bar{g}$ is locally agenda consistent. This proves that $\hat{g}$ satisfies this condition, too.

Finally, for reinforcement, we consider a second vector $v_2 \in \mathbb{Q}^{m!}$ and a scalar $\lambda_2 \in \mathbb{N}$ such that $\lambda_2 v_2 \in \mathbb{Z}^{m!}$. Moreover, we suppose that $\hat{g}(v_1, X) \cap \hat{g}(v_2, X) \neq \emptyset$ for some $X \in \mathcal{F}(A)$. Since $\lambda_1 \lambda_2 (v_1 + v_2) \in \mathbb{Z}^{m!}$, $\bar{g}(\lambda_1 \lambda_2 v_1, X) = \bar{g}(\lambda_1 v_1, X)$, and $\bar{g}(\lambda_1 \lambda_2 v_2, X) = \bar{g}(\lambda_2 v_2, X)$, we derive that $\hat{g}(v_1 + v_2, X) = \bar{g}(\lambda_1 \lambda_2 (v_1 + v_2), X) = \bar{g}(\lambda_1 \lambda_2 v_1, X) \cap \bar{g}(\lambda_1 \lambda_2 v_2, X) = \bar{g}(\lambda_1 v_1, X) \cap \bar{g}(\lambda_2 v_2, X) = \hat{g}(v_1, X) \cap \hat{g}(v_2, X)$. Thus, $\hat{g}$ also satisfies reinforcement. $\square$

Since $f(R, X) = \hat{g}(v(R), X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$, we will from now on analyze the function $\hat{g}$. To this end, we define the sets $D_\triangleright = \{v \in \mathbb{Q}^{m!} \colon \triangleright \in \hat{g}(v, X)\}$ for all feasible sets $X \in \mathcal{F}(A)$ and rankings $\triangleright \in \mathcal{R}(X)$. Intuitively, $D_\triangleright$ is the domain of points $v \in \mathbb{Q}^{m!}$ such that $\triangleright \in \hat{g}(v, X)$. Moreover, we define $\bar{D}_\triangleright$ as the closure of $D_\triangleright$ with respect to $\mathbb{R}^{m!}$. First, note that the sets $\bar{D}_\triangleright$ are symmetric. In more detail, because $\hat{g}$ is neutral, it holds for all permutations $\tau : A \to A$, vectors $v \in \mathbb{Q}^{m!}$, feasible sets $X \in \mathcal{F}(A)$, and rankings $\triangleright \in \mathcal{R}(X)$ that $\tau(\triangleright) \in \hat{g}(\tau(v), \tau(X))$ if $\triangleright \in \hat{g}(v, X)$. By the definition of $\bar{D}_\triangleright$, it hence follows that $\tau(v) \in \bar{D}_{\tau(\triangleright)}$ if $v \in \bar{D}_\triangleright$. Moreover, we note that $\bigcup_{\triangleright \in \mathcal{R}(X)} D_\triangleright = \mathbb{Q}^{m!}$ for every feasible set $X \in \mathcal{F}(A)$ as the domain of $\hat{g}$ is $\mathbb{Q}^{m!} \times \mathcal{F}(A)$. Since there are only finitely many rankings in $\mathcal{R}(X)$, this also implies that $\bigcup_{\triangleright \in \mathcal{R}(X)} \bar{D}_\triangleright = \mathbb{R}^{m!}$ for all feasible sets $X \in \mathcal{F}(A)$. Finally, we note that the sets $D_\triangleright$ are $\mathbb{Q}$-convex: it holds for all $v, v' \in D_\triangleright$ and all $\lambda \in [0, 1] \cap \mathbb{Q}$ that $\lambda v + (1-\lambda)v' \in D_\triangleright$. This follows from the reinforcement of $\hat{g}$ because $v, v' \in D_\triangleright$ implies that $\triangleright \in \hat{g}(v, X) = \hat{g}(\lambda v, X)$ and $\triangleright \in \hat{g}(v', X) = \hat{g}((1-\lambda)v', X)$, and therefore also that $\triangleright \in \hat{g}(\lambda v + (1-\lambda)v', X)$. Moreover, if $D_\triangleright$ is $\mathbb{Q}$-convex, then $\bar{D}_\triangleright$ is convex (see Young (1975)). In fact, these sets are even convex cones because $v \in \bar{D}_\triangleright$ entails that $\lambda v \in \bar{D}_\triangleright$ for every $\lambda > 0$. Finally, we infer that the sets $\bar{D}_\triangleright$ are fully dimensional since $\mathcal{R}(X)$ is finite, all sets $\bar{D}_\triangleright$ with $\triangleright \in \mathcal{R}(X)$ are symmetric to each other, and $\bigcup_{\triangleright \in \mathcal{R}(X)} \bar{D}_\triangleright = \mathbb{R}^{m!}$.

Based on these insights, we next show for all feasible sets $X \in \mathcal{F}(A)$ and rankings $\triangleright, \triangleright' \in \mathcal{R}(X)$ that there is a non-zero vector that *separates* $\bar{D}_\triangleright$ *from* $\bar{D}_{\triangleright'}$. More formally, this means that there is a vector $u^{\triangleright,\triangleright'} \in \mathbb{R}^{m!}$ such that $vu^{\triangleright,\triangleright'} \geq 0$ if $v \in \bar{D}_\triangleright$ and $vu^{\triangleright,\triangleright'} \leq 0$ if $v \in \bar{D}_{\triangleright'}$ (here, vu denotes the standard scalar product).

Lemma 4. *For all feasible sets $X \in \mathcal{F}(A)$ and distinct rankings $\triangleright, \triangleright' \in \mathcal{R}(X)$, there is a non-zero vector $u^{\triangleright,\triangleright'} \in \mathbb{R}^{m!}$ such that $vu^{\triangleright,\triangleright'} \geq 0$ if $v \in \bar{D}_\triangleright$ and $vu^{\triangleright,\triangleright'} \leq 0$ if $v \in \bar{D}_{\triangleright'}$.*

Proof. Consider an arbitrary feasible set $X \in \mathcal{F}(A)$ and two distinct rankings $\triangleright$, $\triangleright' \in \mathcal{R}(X)$. Moreover, let $\bar{D}_\triangleright$ and $\bar{D}_{\triangleright'}$ be defined as explained before the lemma and recall that these sets are convex cones. For finding the non-zero vector $u^{\triangleright,\triangleright'}$ that separates $\bar{D}_\triangleright$ from $\bar{D}_{\triangleright'}$, we aim to apply the separating hyperplane theorem for convex sets (see, e.g., Rockafeller, 1970). This requires us to show that the intersection of the interiors of these two sets are disjoint, i.e., $\text{int}\bar{D}_\triangleright \cap \text{int}\bar{D}_{\triangleright'} = \emptyset$.

Assume for contradiction that this is not the case. Since both $\bar{D}_\triangleright$ and $\bar{D}_{\triangleright'}$ are fully dimensional, this means that there is a point $v \in \text{int}\bar{D}_\triangleright \cap \text{int}\bar{D}_{\triangleright'} \cap \mathbb{Q}^{m!}$. As the first step, we note that $\text{int}\bar{D}_\triangleright$ is a subset of the convex hull of $D_\triangleright$. Because $D_\triangleright$ is $\mathbb{Q}$-convex and $v \in \mathbb{Q}^{m!}$, this entails that $v \in D_\triangleright$. An analogous argument also shows that $v \in D_{\triangleright'}$, so $\triangleright \in \hat{g}(v, X)$ and $\triangleright' \in \hat{g}(v, X)$. Next, there is a profile R such that $f(R, X) = \{\triangleright\}$ because of Lemma 2. By Lemma 3, it follows that $\hat{g}(v', X) = \{\triangleright\}$ for the vector $v' = v(R) \in \mathbb{N}_0^{m!}$. Finally, since $v \in \text{int}\bar{D}_{\triangleright'}$, there is $\lambda \in (0, 1) \cap \mathbb{Q}^{m!}$ such that $v + \lambda v' \in \text{int}\bar{D}_{\triangleright'}$. Using the same reasoning as for v, it follows that $v + \lambda v' \in D_{\triangleright'}$. However, this means that $\triangleright' \in \hat{g}(v + \lambda v', X)$ but $\triangleright' \notin \hat{g}(v, X) \cap \hat{g}(\lambda v', X) = \{\triangleright\}$. This contradicts that $\hat{g}$ is reinforcing, so the assumption that $\text{int}\bar{D}_\triangleright \cap \text{int}\bar{D}_{\triangleright'} \neq \emptyset$ is wrong.

Next, we note that both $\text{int}\bar{D}_\triangleright$ and $\text{int}\bar{D}_{\triangleright'}$ are open (by the definition of the interior), non-empty (because $\bar{D}_\triangleright$ and $\bar{D}_{\triangleright'}$ are fully dimensional) and convex (because $\bar{D}_\triangleright$ and

$\bar{D}_{\rhd'}$ are convex). Hence, we can apply the separating hyperplane theorem for convex sets to infer that there is a non-zero vector $u^{\rhd,\rhd'} \in \mathbb{R}^{m!}$ and a constant $c \in \mathbb{R}$ such that $vu^{\rhd,\rhd'} > c$ for all $v \in \text{int}\bar{D}_{\rhd}$ and $vu^{\rhd,\rhd'} < c$ for all $v \in \text{int}\bar{D}_{\rhd'}$. It thus follows that $vu^{\rhd,\rhd'} \geq c$ for all $v \in \bar{D}_{\rhd}$ and $vu^{\rhd,\rhd'} \leq c$ for all $v \in \bar{D}_{\rhd'}$. Finally, we will prove that $c = 0$ and assume for contradiction that $c > 0$. Now, let $v \in \bar{D}_{\rhd}$, which means that $vu^{\rhd,\rhd'} \geq c$. Moreover, there is a scalar $\lambda > 0$ such that $\lambda vu^{\rhd,\rhd'} < c$ since $c > 0$. However, $\lambda v \in \bar{D}_{\rhd}$ since the sets $\bar{D}_{\rhd}$ are cones. This contradicts that $v'u^{\rhd,\rhd'} \geq c$ for all $v \in \bar{D}_{\rhd}$. Hence, $c \leq 0$ and a symmetric argument rules out that $c < 0$, so c must be 0. This completes the proof of this lemma. $\square$

As the last lemma of this section, we will show that the vectors $u^{\rhd,\rhd'}$ for $\rhd' \in \mathcal{R}(X) \setminus \{\rhd\}$ fully describe the set $\bar{D}_{\rhd}$.

Lemma 5. *Consider a feasible set $X \in \mathcal{F}(A)$ and a ranking $\rhd \in \mathcal{R}(X)$. It holds that $\bar{D}_{\rhd} = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathcal{R}(X) \setminus \{\rhd\} \colon vu^{\rhd,\rhd'} \geq 0\}$, where $u^{\rhd,\rhd'} \in \mathbb{R}^{m!}$ are arbitrary non-zero vectors that separate $\bar{D}_{\rhd}$ from $\bar{D}_{\rhd'}$.*

Proof. Fix an arbitrary feasible set $X \in \mathcal{F}(A)$ and a ranking $\rhd \in \mathcal{R}(X)$. Moreover, we define $S_{\rhd} = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathcal{R}(X) \setminus \{\rhd\} \colon vu^{\rhd,\rhd'} \geq 0\}$ for a simpler notation. First, we note that, by definition, $vu^{\rhd,\rhd'} \geq 0$ for all $v \in \bar{D}_{\rhd}$ and $\rhd' \in \mathcal{R}(X) \setminus \{\rhd\}$. Hence, $v \in \bar{D}_{\rhd}$ implies $v \in S_{\rhd}$, so $\bar{D}_{\rhd} \subseteq S_{\rhd}$.

For the other direction, we first note that $\text{int}S_{\rhd} \neq \emptyset$ since $\text{int}\bar{D}_{\rhd} \neq \emptyset$ and $\bar{D}_{\rhd} \subseteq S_{\rhd}$. Thus, consider a point $v \in \text{int}S_{\rhd}$. This implies that $vu^{\rhd,\rhd'} > 0$ for all $\rhd' \in \mathcal{R}(X) \setminus \{\rhd\}$, which means that $v \notin \bar{D}_{\rhd'}$ for all $\rhd' \in \mathcal{R}(X) \setminus \{\rhd\}$ because $vu^{\rhd,\rhd'} \leq 0$ if $v \in \bar{D}_{\rhd'}$. Since $\bigcup_{\rhd'' \in \mathcal{R}(X)} \bar{D}_{\rhd''} = \mathbb{R}^{m!}$, it follows that $v \in \bar{D}_{\rhd}$. Hence, $\text{int}S_{\rhd} \subseteq \bar{D}_{\rhd}$ and, since $\bar{D}_{\rhd}$ is a closed set, $S_{\rhd} \subseteq \bar{D}_{\rhd}$. This completes the proof of this lemma. $\square$

A.3. Feasible Sets of Size 2

Since the vectors $u^{\rhd,\rhd'}$ completely specify the sets $\bar{D}_{\rhd}$, we will next work towards understanding these vectors in more detail. For this, we first focus on feasible sets of size 2: in this case, we prove that there are very symmetric vectors that separate $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$ and that these symmetric vectors can be described by a bivariate scoring function. We recall for this subsection that we suppose that the preference relations in $\mathcal{R}(A)$ are enumerated and that $\succ^k$ denotes the k-th preference relation in this set.

Lemma 6. *There are non-zero vectors $\hat{u}^{(x,y),(y,x)} \in \mathbb{R}^{m!}$ for all distinct $x, y \in A$ and a bivariate scoring function s such that*

(1) $\hat{u}^{(x,y),(y,x)}$ separates $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$,

(2) $\hat{u}^{(x,y),(y,x)} = -\hat{u}^{(y,x),(x,y)}$,

(3) $\tau(\hat{u}^{(x,y),(y,x)}) = \hat{u}^{(\tau(x),\tau(y)),(\tau(y),\tau(x))}$ for all permutations $\tau : A \to A$, and

(4) $\hat{u}_k^{(x,y),(y,x)} = s(r(\succ^k, x), r(\succ^k, y)) - s(r(\succ^k, y), r(\succ^k, x))$ for all preference relations $\succ^k \in \mathcal{R}(A)$.

Proof. To prove this lemma, we let $u^{(x,y),(y,x)}$ define non-zero vectors that separate $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$ for all distinct $x, y \in A$; such vectors exist due to Lemma 4. We will first show that, up to scaling with a positive scalar, $u^{(x,y),(y,x)}$ is the only vector that separates $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$. Consider for this another non-zero vector $u \neq u^{(x,y),(y,x)}$ that separates $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$. By Lemma 5, we get that $\bar{D}_{(x,y)} = \{v \in \mathbb{R}^{m!} \colon vu^{(x,y),(y,x)} \geq 0\} = \{v \in \mathbb{R}^{m!} \colon vu \geq 0\}$ and $\bar{D}_{(y,x)} = \{v \in \mathbb{R}^{m!} \colon -vu^{(x,y),(y,x)} \geq 0\} = \{v \in \mathbb{R}^{m!} \colon -vu \geq 0\}$. The second equation holds since $-u^{(x,y),(y,x)}$ (resp. $-u$) separates $\bar{D}_{(y,x)}$ from $\bar{D}_{(x,y)}$. This means that $\bar{D}_{(x,y)} \cap \bar{D}_{(y,x)} = \{v \in \mathbb{R}^{m!} \colon vu^{(x,y),(y,x)} = 0\} = \{v \in \mathbb{R}^{m!} \colon vu = 0\}$. Hence, the vectors $u^{(x,y),(y,x)}$ and u are linearly dependent, i.e., there is a non-zero constant $\alpha \in \mathbb{R}$ such that $\alpha u^{(x,y),(y,x)} = u$. Furthermore, $\alpha > 0$ as otherwise $vu^{(x,y),(y,x)} > 0$ for all $v \in \text{int}\bar{D}_{(x,y)}$ but $vu < 0$. Hence, $u^{(x,y),(y,x)}$ is indeed the unique vector that separates $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$ up to multiplication with a positive scalar.

We now define the vectors $\hat{u}^{(x,y),(y,x)}$ by $\hat{u}^{(x,y),(y,x)} = \alpha^{(x,y),(y,x)} u^{(x,y),(y,x)}$ for all pairs of alternatives $x, y \in A$, where $\alpha^{(x,y),(y,x)} > 0$ is a scalar such that $\sum_{k=1}^{m!} |\hat{u}_k^{(x,y),(y,x)}| = 1$ (i.e., the 1-norm of all vectors $\hat{u}^{(x,y),(y,x)}$ is 1). We first note that the vectors $\hat{u}^{(x,y),(y,x)}$ separate $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$ as they are only rescalings of $u^{(x,y),(y,x)}$, so they satisfy Claim (1) of the lemma. Even more, by our previous analysis, the vectors $\hat{u}^{(x,y),(y,x)}$ are the unique vectors that separate $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$ and satisfy $\sum_{k=1}^{m!} |\hat{u}_k^{(x,y),(y,x)}| = 1$.

For the second claim, we note that the vectors $-\hat{u}^{(y,x),(x,y)}$ separate $\bar{D}_{(x,y)}$ from $\bar{D}_{(y,x)}$ and satisfy that $\sum_{k=1}^{m!} | -\hat{u}_k^{(x,y),(y,x)}| = 1$. By the uniqueness of our vectors, we hence derive that $\hat{u}^{(x,y),(y,x)} = -\hat{u}^{(y,x),(x,y)}$, which proves Claim (2) of the lemma.

Next, we observe that for every permutation $\tau : A \to A$, the vector $\tau(\hat{u}^{(x,y),(y,x)})$ separates $\bar{D}_{(\tau(x),\tau(y))}$ from $\bar{D}_{(\tau(y),\tau(x))}$. This follows from the symmetry of the sets $\bar{D}_{(x,y)}$ and $\bar{D}_{(y,x)}$: for every $v \in \bar{D}_{(\tau(x),\tau(y))}$, there is a vector $v' \in \bar{D}_{(x,y)}$ such that $\tau(v') = v$, and it is easy to see that $0 \leq v'\hat{u}^{(x,y),(y,x)} = \tau(v')\tau(\hat{u}^{(x,y),(y,x)}) = v\tau(\hat{u}^{(x,y),(y,x)})$. Since an analogous argument holds for $\bar{D}_{(\tau(y),\tau(x))}$, $\tau(\hat{u}^{(x,y),(y,x)})$ indeed separates $\bar{D}_{(\tau(x),\tau(y))}$ from $\bar{D}_{(\tau(y),\tau(x))}$. Because $\sum_{k=1}^{m!} |\tau(\hat{u}^{(x,y),(y,x)})| = 1$, we derive from the uniqueness of our vectors that $\hat{u}^{(\tau(x),\tau(y)),(\tau(y),\tau(x))} = \tau(\hat{u}^{(x,y),(y,x)})$, which shows Claim (3).

Finally, we will prove Claim (4). To this end, we will construct a bivariate scoring function $\bar{s}$ such that $\hat{u}_k^{(x,y),(y,x)} = \bar{s}(r(\succ^k, x), r(\succ^k, y))$ for all distinct $x, y \in A$ and $\succ^k \in \mathcal{R}(A)$. Based on $\bar{s}$, we then define a second bivariate scoring function s by $s(i,j) = \frac{1}{2}\bar{s}(i,j)$ for all $i, j \in \{1, \dots, m\}$. Since the definition of bivariate scoring functions entails that $\bar{s}(i,j) = -\bar{s}(j,i)$ for all $i, j \in \{1, \dots, m\}$, it follows that $\bar{s}(i,j) = \frac{1}{2}\bar{s}(i,j) - \frac{1}{2}\bar{s}(j,i) = s(i,j) - s(j,i)$ for all $i, j \in \{1, \dots, m\}$. Thus, if $\hat{u}_k^{(x,y),(y,x)} = \bar{s}(r(\succ^k, x), r(\succ^k, y))$ for all distinct $x, y \in A$ and $\succ^k \in \mathcal{R}(A)$, then also $\hat{u}_k^{(x,y),(y,x)} = s(r(\succ^k, x), r(\succ^k, y)) - s(r(\succ^k, y), r(\succ^k, x))$, which completes the proof of Claim (4).

Now, to define the bivariate scoring function $\bar{s}$, we fix a pair of alternatives $x^*, y^* \in A$ and set $\bar{s}(i,j) = \hat{u}_k^{(x^*,y^*),(y^*,x^*)}$ for all $i, j \in \{1, \dots, m\}$, where $\succ^k \in \mathcal{R}(A)$ is an arbitrary preference relation such that $i = r(\succ^k, x^*)$ and $j = r(\succ^k, y^*)$. We first show that $\hat{u}_k^{(x,y),(y,x)} = \bar{s}(r(\succ^k, x), r(\succ^k, y))$ for all $x, y \in A$ and $\succ^k \in \mathcal{R}(A)$. To this end, we consider two alternatives $x, y \in A$ and a preference relation $\succ^k$. Moreover, let $\succ^{k'}$ denote the preference relation used to define $\bar{s}(i,j)$ for $i = r(\succ^k, x)$ and $j = r(\succ^k, y)$, which means that $r(\succ^k, x) = r(\succ^{k'}, x^*)$ and $r(\succ^k, y) = r(\succ^{k'}, y^*)$. Consequently, there is a permutation $\tau : A \to A$ such that $\tau(x^*) = x$, $\tau(y^*) = y$, and $\tau(\succ^{k'}) = \succ^k$. By Claim (3), it holds that $\hat{u}^{(x,y),(y,x)} = \tau(\hat{u}^{(x^*,y^*),(y^*,x^*)})$. In particular, $\hat{u}_k^{(x,y),(y,x)} = \tau(\hat{u}^{(x^*,y^*),(y^*,x^*)})_k = \hat{u}_{k'}^{(x^*,y^*),(y^*,x^*)}$ because $\tau(\succ^{k'}) = \succ^k$. Hence, $\hat{u}_k^{(x,y),(y,x)} = \hat{u}_{k'}^{(x^*,y^*),(y^*,x^*)} = \bar{s}(r(\succ^{k'}, x^*), r(\succ^{k'}, y^*)) = \bar{s}(r(\succ^k, x), r(\succ^k, y))$ for all distinct alternatives $x, y \in A$ and preference relations $\succ^k \in \mathcal{R}(A)$.

It remains to show that $\bar{s}$ is indeed a bivariate scoring function. To this end, we note that $\bar{s}(i,j) = -\bar{s}(j,i)$ because $\bar{s}(r(\succ^k, x), r(\succ^k, y)) = \hat{u}_k^{(x,y),(y,x)} = -\hat{u}_k^{(y,x),(x,y)} = -\bar{s}(r(\succ^k, y), r(\succ^k, x))$ for all $x, y \in A$ and $\succ^k \in \mathcal{R}(A)$. Furthermore, the faithfulness of f entails that $\bar{s}(i,j) \geq 0$ if $i < j$. To make this more precise, fix two distinct alternatives $x, y \in A$ and consider a preference relation $\succ^k \in \mathcal{R}(A)$ with $x \succ y$. By faithfulness, it holds that $(x,y) \in f(\succ^k, \{x,y\}) = \hat{g}(v(\succ^k), \{x,y\})$, so $v(\succ^k) \in \bar{D}_{(x,y)}$. This means that $v(\succ^k)\hat{u}^{(x,y),(y,x)} = \hat{u}_k^{(x,y),(y,x)} \geq 0$, so $\bar{s}(i,j) \geq 0$ if $i < j$. Finally, since $\hat{u}^{(x,y),(y,x)}$ is non-zero, our observations also imply the existence of indices $i, j \in \{1, \dots, m\}$ with $i < j$ and $\bar{s}(i,j) > 0$. ◻

From now on, we will only work with the vectors $\hat{u}^{(x,y),(y,x)}$ constructed in Lemma 6 as they are highly symmetric, and thus denote these vectors by $u^{(x,y),(y,x)}$ (without the hat). Furthermore, we note that Claim (4) of this lemma gives an intuitive interpretation of these vectors: the term $v(R)u^{(x,y),(y,x)}$ effectively computes the score difference between the rankings (x,y) and (y,x) for the profile R. Based on this interpretation, we can, for instance, derive the vectors $u^{(x,y),(y,x)}$ for the plurality rule (up to scaling with a positive scalar): $u_k^{(x,y),(y,x)} = 1$ if x is top-ranked in the preference relation $\succ^k$, $u_k^{(x,y),(y,x)} = -1$ if y is top-ranked in $\succ^k$, and $u_k^{(x,y),(y,x)} = 0$ otherwise. As a second example, it can be checked that Kemeny's rule can be represented by $u_k^{(x,y),(y,x)} = 1$ if $r(\succ^k, x) < r(\succ^k, y)$ and $u_k^{(x,y),(y,x)} = -1$ otherwise. That is, the expression $v(R)u^{(x,y),(y,x)}$ computes for Kemeny's rule the majority margin between x and y with respect to R.

To facilitate the analysis of larger feasible sets, we will next investigate the linear dependence of the vectors $u^{(x,y),(y,x)}$ constructed in Lemma 6.

Lemma 7. *Consider a feasible set $X = \{a_1, \dots, a_\ell\} \in \mathcal{F}(A)$ with $\ell \geq 3$ and the sets $U_1^X = \{u^{(a_1,a_i),(a_i,a_1)} \colon i \in \{2, \dots, \ell\}\}$ and $U_2^X = \{u^{(a_i,a_j),(a_j,a_i)} \colon i, j \in \{1, \dots, \ell\} \colon i < j\}$. The following claims are true:*

(1) The set U_1^X is linearly independent.

(2) If U_2^X is linearly dependent, then $u^{(a_i,a_j),(a_j,a_i)} = u^{(a_i,a_1),(a_1,a_i)} + u^{(a_1,a_j),(a_j,a_1)}$ for all $i, j \in \{2, \dots, \ell\}$ with $i < j$.

Proof. Let $X \in \mathcal{F}(A)$ denote an arbitrary feasible set with at least three alternatives and define U_1^X and U_2^X as in the lemma. Before presenting the proofs of the two claims of this lemma, we will discuss an auxiliary statement that relates the profiles R^x of Lemma 1 with $f(R^x, X) = \{\rhd \in \mathcal{R}(X) \colon \forall y \in X \setminus \{x\} \colon x \rhd y\}$ to the vectors $u^{(y,z),(z,y)}$. Based on this auxiliary claim, we then prove the lemma.

Auxiliary Claim: $v(R^x)u^{(y,z),(z,y)} = 0$ **and** $v(R^x)u^{(x,y),(y,x)} = v(R^x)u^{(x,z),(x,z)} > 0$ **for all distinct** $x, y, z \in X$**.**

To prove this auxiliary claim, we fix a triple of alternatives $x, y, z \in X$ and let R^x denote the profile constructed in Lemma 1 with $f(R^x, X) = \{\rhd \in \mathcal{R}(X) \colon \forall y \in X \setminus \{x\} \colon x \rhd y\}$. First, we will show that $v(R^x)u^{(y,z),(z,y)} = 0$ and note for this that there are two rankings $\rhd', \rhd'' \in f(R^x, X)$ such that y, z are adjacent in both $\rhd'$ and $\rhd''$ but $y \rhd' z$ and $z \rhd'' y$. By local agenda consistency, it thus follows that $f(R^x, \{y, z\}) = \hat{g}(v(R^x), \{y, z\}) = \{(y, z), (z, y)\}$. This implies that $v(R^x) \in \bar{D}_{(y,z)} \cap \bar{D}_{(z,y)}$. The definition of $u^{(y,z),(z,y)}$ hence requires that $v(R^x)u^{(y,z),(z,y)} \geq 0$ as $v(R^x) \in \bar{D}_{(y,z)}$ and $v(R^x)u^{(y,z),(z,y)} \leq 0$ as $v(R^x) \in \bar{D}_{(z,y)}$, so $v(R^x)u^{(y,z),(z,y)} = 0$.

Next, we will show that $v(R^x)u^{(x,y),(y,x)} = v(R^x)u^{(x,z),(z,x)} > 0$. To this end, we let τ denote the permutation with $\tau(y) = z$, $\tau(z) = y$, and $\tau(w) = w$ for all $w \in A \setminus \{y, z\}$. By the symmetry of R^x (see the proof of Lemma 1), it holds that $\tau(R^x) = R^x$ (up to renaming the voters). In more detail, we recall here that R^x is defined by $R^x = \sum_{w \in X \setminus \{x\}} R^{x,w}$, where the profiles $R^{x,w}$ consists of $\frac{m!}{2}$ voters such that each preference relation $\succ \in \mathcal{R}(A)$ with $x \succ w$ is reported by one voter. It can be checked that, except for reordering the voters, $\tau(R^{x,y}) = R^{x,z}$, $\tau(R^{x,z}) = R^{x,y}$, and $\tau(R^{x,w}) = R^{x,w}$ for all $w \in X \setminus \{x, y, z\}$. Hence, $\tau(R^x) = \sum_{w \in X \setminus \{x\}} \tau(R^{x,w}) = \sum_{w \in X \setminus \{x\}} R^{x,w} = R^x$. We therefore conclude that $v(R^x)u^{(x,y),(y,x)} = \tau(v(R^x))\tau(u^{(x,y),(y,x)}) = v(\tau(R^x))u^{(\tau(x),\tau(y)),(\tau(y),\tau(x))} = v(R^x)u^{(x,z),(z,x)}$. Due to these computations, it suffices to show that $v(R^x)u^{(x,y),(y,x)} > 0$ to prove the second part of our auxiliary claim.

To do so, we first recall that $f(R^x, \{x, y\}) = \{(x, y)\}$ (see Lemma 1 for details). In particular, this implies that $v(R^x) \in \bar{D}_{(x,y)}$ and thus that $v(R^x)u^{(x,y),(y,x)} \geq 0$. We next suppose for contradiction that $v(R^x)u^{(x,y),(y,x)} = 0$. In this case, let k denote an index such that $u^{(x,y),(y,x)}_k < 0$. To see that such an index exist, let τ denote the permutation with $\tau(x) = y$, $\tau(y) = x$, and $\tau(w) = w$ for all $w \in A \setminus \{x, y\}$. Lemma 6 entails that $u^{(x,y),(y,x)} = -u^{(y,x),(x,y)} = -u^{(\tau(x),\tau(y)),(\tau(y),\tau(x))} = -\tau(u^{(x,y),(y,x)})$. Since $u^{(x,y),(y,x)}$ is a non-zero vector, this shows that there is an index k with $u^{(x,y),(y,x)}_k < 0$. Hence, it follows that $v(\lambda R^x + \succ^k)u^{(x,y),(y,x)} = u^{(x,y),(y,x)}_k < 0$ for every $\lambda \in \mathbb{N}$. By the definition of $u^{(x,y),(y,x)}$, this means that $v(\lambda R^x + \succ^k) \notin \bar{D}_{(x,y)}$ for every $\lambda \in \mathbb{N}$. However, the continuity of f implies that there is $\lambda' \in \mathbb{N}$ such that $f(\lambda' R^x + \succ^k, \{x, y\}) = \hat{g}(v(\lambda' R^x + \succ^k), \{x, y\}) = \{(x, y)\}$. This means that $v(\lambda' R^x + \succ^k) \in \bar{D}_{(x,y)}$, which contradicts our previous observation. Hence, the assumption that $v(R^x)u^{(x,y),(y,x)} = 0$ must be wrong, and we finally conclude that $v(R^x)u^{(x,y),(y,x)} = v(R^x)u^{(x,z),(z,x)} > 0$.

Claim (1): The set U_1^X is linearly independent.

To prove this claim, we let a_i denote an arbitrary alternative in $X \setminus \{a_1\}$ and we will show that $u^{(a_1,a_i),(a_i,a_1)}$ is linearly independent of the other vectors in U_1^X. To this end, let R^{a_i} denote the profile constructed in Lemma 1 with $f(R^{a_i}, X) = \{\rhd \in \mathcal{R}(X) \colon \forall x \in X \setminus \{a_i\} \colon a_i \rhd x\}$. By our auxiliary claim and the fact that $u^{(a_1,a_i),(a_i,a_1)} = -u^{(a_i,a_1),(a_1,a_i)}$, it follows that $v(R^{a_i})u^{(a_1,a_i),(a_i,a_1)} = -v(R^{a_i})u^{(a_i,a_1),(a_1,a_i)} < 0$ and $v(R^{a_i})u^{(a_1,a_j),(a_j,a_1)} = 0$ for all $a_j \in X \setminus \{a_1, a_i\}$. This implies that $u^{(a_1,a_i),(a_i,a_1)}$ is linearly independent of the other vectors $u^{(a_1,a_j),(a_j,a_1)} \in U_1^X \setminus \{u^{(a_1,a_i),(a_i,a_1)}\}$, so U_1^X is linearly independent.

Claim (2): If U_2^X is linearly dependent, then $u^{(a_i,a_j),(a_j,a_i)} = u^{(a_i,a_1),(a_1,a_i)} + u^{(a_1,a_j),(a_j,a_1)}$ for all $i,j \in \{2,\dots,\ell\}$ with $i<j$.

Assume that the set U_2^X is linearly dependent. As a warm-up, we first prove the claim for the case that $|X|=3$ and therefore assume that $X=\{a_1,a_2,a_3\}$. By Claim (1), we have that U_1^X is linearly independent, so there are values λ_2, $\lambda_3 \in \mathbb{R}$ such that $u^{(a_2,a_3),(a_3,a_2)} = \lambda_2 u^{(a_1,a_2),(a_2,a_1)} + \lambda_3 u^{(a_1,a_3),(a_3,a_1)}$. Now, consider the profile R^{a_1} of Lemma 1, for which our auxiliary claim shows that $v(R^{a_1})u^{(a_2,a_3),(a_3,a_2)} = 0$ and $v(R^{a_1})u^{(a_1,a_2),(a_2,a_1)} = v(R^{a_1})u^{(a_1,a_3),(a_3,a_1)} > 0$. We derive that $0 = v(R^{a_1})u^{(a_2,a_3),(a_3,a_2)} = v(R^{a_1})(\lambda_2 u^{(a_1,a_2),(a_2,a_1)} + \lambda_3 u^{(a_1,a_3),(a_3,a_1)}) = (\lambda_2+\lambda_3)v(R^{a_1})u^{(a_1,a_2),(a_2,a_1)}$ and hence conclude that $\lambda_2 = -\lambda_3$. Finally, we consider the profile R^{a_2}, for which the auxiliary claim shows that $v(R^{a_2})u^{(a_2,a_3),(a_3,a_2)} = v(R^{a_2})u^{(a_2,a_1),(a_1,a_2)} > 0$ and $v(R^{a_2})u^{(a_1,a_3),(a_3,a_1)} = 0$. Since $u^{(a_2,a_1),(a_1,a_2)} = -u^{(a_1,a_2),(a_2,a_1)}$, we infer that $\lambda_2=-1$ and therefore $\lambda_3 = 1$. This means that $u^{(a_2,a_3),(a_3,a_2)} = -u^{(a_1,a_2),(a_2,a_1)} + u^{(a_1,a_3),(a_3,a_1)} = u^{(a_2,a_1),(a_1,a_2)} + u^{(a_1,a_3),(a_3,a_1)}$.

Next, we suppose that $|X|\geq 4$ and that $u^{(x,y),(y,x)} \in U_2^X$ is a vector that is linearly depending on the other vectors in U_2^X. Furthermore, let $U_2^- = U_2^X \setminus \{u^{(x,y),(y,x)}\}$. By the linear dependence, there are values $\lambda_{(x',y'),(y',x')} \in \mathbb{R}$ such that $u^{(x,y),(y,x)} = \sum_{u^{(x',y'),(y',x')}\in U_2^-} \lambda_{(x',y'),(y',x')} u^{(x',y'),(y',x')}$. Moreover, we define $\mathcal{T}_1 = \{\tau \in A^A\colon \tau(x) = x \land \tau(y) = y\}$ as the set of permutations that map x to x and y to y. By Lemma 6, it holds that $u^{(x,y),(y,x)} = u^{(\tau(x),\tau(y)),(\tau(y),\tau(x))} = \tau(u^{(x,y),(y,x)})$ for all or all permutations $\tau \in \mathcal{T}_1$ and therefore also that $u^{(x,y),(y,x)} = \frac{1}{(m-2)!}\sum_{\tau\in\mathcal{T}_1}\tau(u^{(x,y),(y,x)}) = \frac{1}{(m-2)!}\sum_{\tau\in\mathcal{T}_1}\sum_{u^{(x',y'),(y',x')}\in U_2^-}\lambda_{(x',y'),(y',x')}\tau(u^{(x',y'),(y',x')})$.

As our next step, we will simplify this representation and hence define $\lambda'_{(v,w),(w,v)} = \frac{1}{(m-2)!}\sum_{x',y'\in X\colon x'\neq y'}\sum_{\tau\in\mathcal{T}_1\colon \tau(x')=v\land\tau(y')=w}\lambda_{(x',y'),(y',x')}$ for all distinct alternatives $v,w\in A$. It then follows that $u^{(x,y),(y,x)} = \sum_{x',y'\in A\colon x'\neq y'}\lambda'_{(x',y'),(y',x')}u^{(x',y'),(y',x')}$. Now, we first note that $\lambda'_{(x,y),(y,x)} = \lambda'_{(y,x),(x,y)} = 0$ because $\tau(x)=x$ and $\tau(y)=y$ for all $\tau\in\mathcal{T}_1$ and $u^{(x,y),(y,x)}, u^{(y,x),(x,y)} \notin U_2^-$. Next, we consider two pairs of alternatives $v,w\in A\setminus\{x,y\}$ and $x',y'\in X\setminus\{x,y\}$. There are precisely $(m-4)!$ permutations in $\mathcal{T}_1$ with $\tau(x')=v$ and $\tau(y')=w$ and another $(m-4)!$ permutations with $\tau(x')=w$ and $\tau(y')=v$. Hence, $\lambda'_{(v,w),(w,v)} = \lambda'_{(w,v),(v,w)}$. Since Lemma 6 shows that $u^{(v,w),(w,v)} = -u^{(w,v),(v,w)}$, we can cancel out the terms that only concern alternatives $v,w\in A\setminus\{x,y\}$. More formally, this means that $u^{(x,y),(y,x)} = \sum_{z\in A\setminus\{x,y\}}\lambda'_{(x,z),(z,x)}u^{(x,z),(z,x)} + \lambda'_{(z,y),(y,z)}u^{(z,y),(y,z)}$. (Note here that we can always replace $u^{(z,x),(x,z)}$ with $-u^{(x,z),(z,x)}$ and push the minus into $\lambda'_{(x,z),(z,x)}$ if required).

Furthermore, for all $z,z'\in A\setminus\{x,y\}$, there are precisely $(m-3)!$ permutations with $\tau(x)=x$, $\tau(y)=y$, and $\tau(z')=z$. Hence, each vector $u^{(x,z'),(z',x)}$ is mapped to $u^{(x,z),(z,x)}$ equally often. This implies that $\lambda'_{(x,z),(z,x)} = \lambda'_{(x,z'),(z',x)}$ for all $z,z'\in A\setminus\{x,y\}$. An analogous argument also shows that $\lambda'_{(z,y),(y,z)} = \lambda'_{(z',y),(y,z')}$ for all $z,z'\in A\setminus\{x,y\}$. Finally, by considering again the profiles R^z of Lemma 1, we derive that $\lambda'_{(x,z),(z,x)} = \lambda'_{(z,y),(y,z)}$ for all $z\in A\setminus\{x,y\}$ since our auxiliary claim shows that $v(R^z)u^{(x,y),(y,x)} = v(R^z)u^{(x,w),(w,x)} = v(R^z)u^{(w,y),(y,w)} = 0$ for all $w\in X\setminus\{x,y,z\}$, and

$v(R^z)u^{(x,z),(z,x)} = -v(R^z)u^{(z,y),(y,z)} < 0$. Hence, there is $\lambda \neq 0$ such that $u^{(x,y),(y,x)} = \lambda \sum_{z\in A\setminus\{x,y\}} u^{(x,z),(z,x)} + u^{(z,y),(y,z)}$.

As our next step, we will derive a representation of $u^{(x,y),(y,x)}$ based on less "intermediate" alternatives. For this, let B denote the set of alternatives z such that the vectors $u^{(x,z),(z,x)}$ and $u^{(z,y),(y,z)}$ still appear in the presentation of $u^{(x,y),(y,x)}$. In particular, $B = A \setminus \{x, y\}$ in the beginning. We will now give a construction that allows to remove one alternative from B if $|B| > 1$. Thus, choose an arbitrary alternative $z \in B\setminus\{a_1\}$ and consider the permutations τ_1 and τ_2 that only swap x and z, and y and z, respectively (i.e., $\tau_1(x) = z$, $\tau_1(z) = x$, and $\tau_1(w) = w$ for all $w \in A \setminus \{x, z\}$; $\tau_2(y) = z$, $\tau_2(z) = y$, and $\tau_2(w) = w$ for all $w \in A \setminus \{y, z\}$). Claim (3) of Lemma 6 shows that

$$\begin{aligned}u^{(z,y),(y,z)} &= \tau_1(u^{(x,y),(y,x)}) = \lambda \sum_{w\in B} \tau_1(u^{(x,w),(w,x)}) + \tau_1(u^{(w,y),(y,w)})\\ &= \lambda u^{(z,x),(x,z)} + \lambda u^{(x,y),(y,x)} + \lambda \sum_{w\in B\setminus\{z\}} u^{(z,w),(w,z)} + u^{(w,y),(y,w)}\end{aligned}$$

$$\begin{aligned}u^{(x,z),(z,x)} &= \tau_2(u^{(x,y),(y,x)}) = \lambda \sum_{w\in B} \tau_2(u^{(x,w),(w,x)}) + \tau_2(u^{(w,y),(y,w)})\\ &= \lambda u^{(x,y),(y,x)} + \lambda u^{(y,z),(z,y)} + \lambda \sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)} + u^{(w,z),(z,w)}.\end{aligned}$$

Since $\sum_{w\in B\setminus\{z\}} u^{(z,w),(w,z)} = -\sum_{w\in B\setminus\{z\}} u^{(w,z),(z,w)}$, we infer the following equation by summing up our last two equalities.

$$\begin{aligned}u^{(x,z),(z,x)} + u^{(z,y),(y,z)} =&\lambda u^{(z,x),(x,z)} + \lambda u^{(y,z),(z,y)} + 2\lambda u^{(x,y),(y,x)}\\ &+ \lambda \sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)} + u^{(w,y),(y,w)}\end{aligned}$$

Next, by dividing by 2λ and substituting $u^{(x,y),(y,x)}$ with $\lambda \sum_{w\in B} u^{(x,w),(w,x)} + u^{(w,y),(y,w)}$, we get that

$$\begin{aligned}\frac{1}{2\lambda}u^{(x,z),(z,x)} + \frac{1}{2\lambda}u^{(z,y),(y,z)} =&(\frac{1}{2} - \lambda)(u^{(z,x),(x,z)} + u^{(y,z),(z,y)})\\ &+ (\frac{1}{2} + \lambda) \sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)} + u^{(w,y),(y,w)}.\end{aligned}$$

This simplifies to $(\frac{1}{2\lambda} - \lambda + \frac{1}{2})(u^{(x,z),(z,x)} + u^{(z,y),(y,z)}) = (\frac{1}{2} + \lambda) \sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)} + u^{(w,y),(y,w)}$. Now, first note that Claim (1) shows that $u^{(x,z),(z,x)} + u^{(z,y),(y,z)} \neq 0$ as these two vectors are linearly independent. Furthermore, if $\sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)} + u^{(w,y),(y,w)} = 0$, then $u^{(x,y),(y,x)} = \lambda(u^{(x,z),(z,x)} + u^{(z,y),(y,z)})$ since $u^{(x,y),(y,x)} = \lambda \sum_{w\in B} u^{(x,w),(w,x)} + u^{(w,y),(w,x)}$. By applying the permutation τ that only swaps a_1 and z, we equivalently get that $u^{(x,y),(y,x)} = \lambda(u^{(x,a_1),(a_1,x)} + u^{(a_1,y),(y,a_1)})$ and we can then proceed with the steps in the subsequent paragraph. Hence, we assume that $\sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)}+u^{(w,y),(y,w)} \neq 0$. For this case, we note that $(\frac{1}{2\lambda}-\lambda+\frac{1}{2}) = 0$ is only true if $\lambda = -\frac{1}{2}$ or $\lambda = 1$. Now, if $\lambda = 1$, then $0 = (\frac{1}{2}+1)\sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)}+u^{(w,y),(y,w)}$, but this contradicts that $\sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)} + u^{(w,y),(y,w)} \neq 0$. On the other hand, $\lambda \neq -\frac{1}{2}$ as otherwise $v(R^x)u^{(x,y),(y,x)} > 0$ and $\lambda v(R^x)\sum_{w\in B} u^{(x,w),(w,x)} + u^{(w,y),(y,w)} < 0$ for the profile R^x with $f(R^x, A) = \{\rhd \in \mathcal{R}(A)\colon \forall y \in X\setminus\{x\}\colon x \rhd y\}$. This follows from our auxiliary claim as $v(R^x)u^{(x,w),(w,x)} > 0$ for every $w \in A\setminus\{x\}$ and $v(R^x)u^{(w,y),(y,w)} = 0$ for all $w \in A\setminus\{x,y\}$. Hence, we can conclude that

$$u^{(x,z),(z,x)} + u^{(z,y),(y,z)} = \frac{\frac{1}{2}+\lambda}{\frac{1}{2\lambda}-\lambda+\frac{1}{2}} \sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)} + u^{(w,y),(y,w)}.$$

By setting $\lambda' = \lambda + \lambda\frac{\frac{1}{2}+\lambda}{\frac{1}{2\lambda}-\lambda+\frac{1}{2}}$, we then get that $u^{(x,y),(y,x)} = \lambda'\sum_{w\in B\setminus\{z\}} u^{(x,w),(w,x)} + u^{(w,y),(y,w)}$, and we have thus removed z from our set B. We can clearly repeat this until $B = \{a_1\}$, so $u^{(x,y),(y,x)} = \lambda(u^{(x,a_1),(a_1,x)} + u^{(a_1,y),(y,a_1)})$ for some $\lambda \in \mathbb{R}$. Moreover, $\lambda \neq 0$ as $u^{(x,y),(y,x)}$ is a non-zero vector.

As the last step, we need to show that $\lambda = 1$. To this end, we consider again the profile R^x of Lemma 1: it holds that $v(R^x)u^{(x,y),(y,x)} = v(R^x)u^{(x,a_1),(a_1,x)} > 0$ and $v(R^x)u^{(a_1,y),(y,a_1)} = 0$ due to our auxiliary claim. So, it follows that $\lambda = 1$. Finally, we consider two arbitrary alternatives a_i, a_j with $i, j \in \{2,\dots,\ell\}$ and $i < j$ and let τ denote a permutation with $\tau(x) = a_i$, $\tau(y) = a_j$, and $\tau(a_1) = a_1$. It follows from Lemma 6 that $u^{(a_i,a_j),(a_j,a_i)} = \tau(u^{(x,y),(y,x)}) = \tau(u^{(x,a_1),(a_1,x)}) + \tau(u^{(a_1,y),(y,a_1)}) = u^{(a_i,a_1),(a_1,a_i)} + u^{(a_1,a_j),(a_1,a_j)}$, so the lemma is proven. $\square$

For an example of Lemma 7, we turn back to the plurality rule and Kemeny's rule and recall that the vectors $u^{(x,y),(y,x)}$ have been defined after Lemma 6 for these rules. In particular, for the plurality rule, the set U_2^X is linearly dependent if $|X| \geq 3$, which can be verified by considering any three alternatives $x, y, z \in X$: if x is top-ranked in $\succ^k$, then $u_k^{(y,z),(z,y)} = 0 = -1 + 1 = u_k^{(y,x),(x,y)} + u_k^{(x,z),(z,x)}$; if y is top-ranked in $\succ^k$, then $u_k^{(y,z),(z,y)} = 1 = 1 + 0 = u_k^{(y,x),(x,y)} + u_k^{(x,z),(z,x)}$; if z is top-ranked in $\succ^k$, then $u_k^{(y,z),(z,y)} = -1 = 0 - 1 = u_k^{(y,x),(x,y)} + u_k^{(x,z),(z,x)}$; finally, if some other alternative is top-ranked in $\succ^k$, then $u_k^{(y,z),(z,y)} = 0 = u_k^{(y,x),(x,y)} + u_k^{(x,z),(z,x)}$. By contrast, for Kemeny's rule, it can be shown that the set U_2^X is linearly independent. To this end, it suffices to consider two preference relations $\succ^{k^1}$ and $\succ^{k^2}$ that only differ on the order of two alternatives x and y, i.e., $x \succ^{k^1} y$, $y \succ^{k^2} x$, and for all $w, z \in A$ with $\{w,z\} \neq \{x,y\}$, it holds that $w \succ^{k^1} z$ if and only if $w \succ^{k^2} z$. For these preference relations, it holds that $u_{k^1}^{(x,y),(y,x)} = 1$, $u_{k^2}^{(x,y),(y,x)} = -1$, and $u_{k^1}^{(w,z),(z,w)} = u_{k^2}^{(w,z),(z,w)}$ for all $w, z \in A$ with $\{w,z\} \neq \{x,y\}$. Hence, the vector $u^{(x,y),(y,x)}$ is linearly independent of $U_2^X \setminus \{u^{(x,y),(y,x)}\}$ and the set U_2^X is thus linearly independent. Alternatively, one can infer this insight also by using that $vu^{(x,y),(y,x)}$ computes the majority margin between x and y for Kemeny's rule as the majority margins of all pairs of alternatives are well-known to be linearly independent. More generally, it can be shown that the set U_2^X is linearly dependent if the vectors in U_2^X correspond to a positional scoring rule and linearly independent if they correspond to a bivariate scoring rule that is not a positional scoring rule.

A.4. Feasible Sets of Size Greater 2

We now turn to feasible sets X with $|X| > 2$ and, as a first step, relate the vectors $u^{\rhd,\rhd'}$ for rankings $\rhd, \rhd'$ on large feasible sets with those for smaller feasible sets. To this end, recall that $C(X,Y) = \{\succ \in \mathcal{R}(Y) \colon \forall x, y \in X, z \in Y \setminus X \colon x \succ z \text{ if and only if } y \succ z\}$ denotes the set of rankings on Y in which the alternatives in X appear consecutively.

Lemma 8. *Consider two feasible set $X, Y \in \mathcal{F}(A)$ with $X \subseteq Y$ and two rankings $\rhd, \rhd' \in \mathcal{R}(Y) \cap C(X,Y)$ such that $\rhd|_X \neq \rhd'|_X$. The vector $u^{\rhd|_X,\rhd'|_X}$ separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$.*

Proof. Consider two feasible sets $X, Y \in \mathcal{F}(A)$ with $X \subseteq Y$ and two rankings $\rhd, \rhd' \in \mathcal{R}(Y)$ that satisfy the conditions of the lemma. The central insight for this lemma is that $\hat{g}$ is locally agenda consistent (see Lemma 3). Thus, if $\rhd \in \hat{g}(v, Y) \cap C(X,Y)$ for some $v \in \mathbb{Q}^{m!}$, then $\rhd|_X \in \hat{g}(v, X)$. Consequently, $v \in D_\rhd$ implies that $v \in D_{\rhd|_X}$, which equivalently means that $D_\rhd \subseteq D_{\rhd|_X}$. From this, we infer that $\bar{D}_\rhd \subseteq \bar{D}_{\rhd|_X}$ and a symmetric reasoning shows that $\bar{D}_{\rhd'} \subseteq \bar{D}_{\rhd'|_X}$. Now, consider a non-zero vector $u^{\rhd|_X,\rhd'|_X}$ that satisfies that $vu^{\rhd|_X,\rhd'|_X} \geq 0$ if $v \in \bar{D}_{\rhd|_X}$ and $vu^{\rhd|_X,\rhd'|_X} \leq 0$ if $v \in \bar{D}_{\rhd'|_X}$. Since $\bar{D}_\rhd \subseteq \bar{D}_{\rhd|_X}$ and $\bar{D}_{\rhd'} \subseteq \bar{D}_{\rhd'|_X}$, this vector clearly separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$, too. $\square$

We note that Lemma 8 is rather general, but we will only use it for a special case: given a feasible set X with $|X| \geq 3$ and two rankings $\rhd, \rhd' \in \mathcal{R}(X)$ such that $\rhd \setminus \rhd' = \{(x,y)\}$ for some pair of alternatives $x, y \in X$, Lemma 8 shows that $u^{(x,y),(y,x)}$ separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$. We thus assume from now that $u^{\rhd,\rhd'} = u^{(x,y),(y,x)}$ for all rankings $\rhd, \rhd' \in \mathcal{R}(X)$ with $\rhd \setminus \rhd' = \{(x,y)\}$ for some pair of alternatives $x, y \in X$. This allows us to transfer the insights of Appendix A.3 to larger feasible sets.

We next aim to find a representation of the vectors $u^{\rhd,\rhd'}$ for rankings $\rhd, \rhd'$ with $|\rhd \setminus \rhd'| > 1$. To this end, we will show that the vector $u = \sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$ for all feasible sets $X \in \mathcal{F}(A)$ and distinct rankings $\rhd, \rhd' \in \mathcal{R}(X)$. For doing so, we will heavily rely on Lemma 7; we thus recall that $U_1^X = \{u^{(a_1,a_i),(a_i,a_1)} \colon i \in \{2, \dots, \ell\}\}$ and $U_2^X = \{u^{(a_i,a_j),(a_j,a_i)} \colon i, j \in \{1, \dots, \ell\} \colon i < j\}$ for an arbitrary feasible set $X = \{a_1, \dots, a_\ell\}$. Now, Lemma 7 shows that either the set U_2^X is linearly independent (as, e.g., for Kemeny's rule), or every strict superset of U_1^X is linearly dependent (as, e.g., for the plurality rule). We consider these cases separately and focus first on the case that U_1^X is a maximal linearly independent set.

Lemma 9. *Consider a feasible set $X \in \mathcal{F}(A)$ with $|X| \geq 3$ and two distinct rankings $\rhd, \rhd' \in \mathcal{R}(X)$ and suppose that the set U_2^X is linearly dependent. The vector $u = \sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ is non-zero and separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$.*

Proof. Consider a feasible set $X = \{a_1, \dots, a_\ell\} \in \mathcal{F}(A)$ such that $|X| \geq 3$ and U_2^X is linearly dependent, and let $\rhd = a_1, \dots, a_\ell$ and $\rhd' = a_1', \dots, a_\ell'$ denote two distinct rankings in $\mathcal{R}(X)$. Moreover, we define $u = \sum_{(x,y)\in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ and first show that u is a non-zero vector. To this end, let i denote the smallest index such that $a_i \neq a'_i$. This implies that $(a_i, a_{i+1}) \in \rhd \setminus \rhd'$ and $(a_j, a_i) \notin \rhd \setminus \rhd'$ for all $j \in \{1, \dots, m\}$. Next, let $v(R^{a_i})$ denote the profile constructed in Lemma 1 with $f(R^{a_i}, X) = \{\rhd \in \mathcal{R}(X) \colon \forall x \in X \setminus \{a_i\} \colon a_i \rhd x\}$. By the auxiliary claim in the proof of Lemma 7, it holds $v(R^{a_i})u^{(x,y),(y,x)} = 0$ for all $x, y \in X \setminus \{a_1\}$ and $v(R^{a_i})u^{(a_i,y),(y,a_i)} > 0$ for all $y \in X \setminus \{a_i\}$. Combining our observations then shows that $v(R^{a_i})u = v(R^{a_i}) \sum_{(x,y)\in \rhd \setminus \rhd'} u^{(x,y),(y,x)} = v(R^{a_i}) \sum_{y \in X \colon (a_i,y)\in \rhd \setminus \rhd'} u^{(a_i,y),(y,a_i)} > 0$, which implies that u is a non-zero vector.

Next, we will show that u indeed separates $\bar{D}_{\rhd}$ from $\bar{D}_{\rhd'}$. To this end, let v denote an arbitrary vector in $\bar{D}_{\rhd}$. Due to the local agenda consistency of $\hat{g}$, it holds that $\bar{D}_{\rhd} \subseteq \bar{D}_{(a_i,a_{i+1})}$ for all $i \in \{1, \dots, \ell - 1\}$, so $v \in \bar{D}_{(a_i,a_{i+1})}$ for all these sets. This means that $vu^{(a_i,a_{i+1}),(a_{i+1},a_i)} \geq 0$ for all $i \in \{1, \dots, \ell - 1\}$. By Lemma 7 and the linear dependence of U_2^X, this is equivalent to $vu^{(a_i,a_1),(a_1,a_i)} + vu^{(a_1,a_{i+1}),(a_{i+1},a_1)} \geq 0$. Since $u^{(a_i,a_1),(a_1,a_i)} = -u^{(a_1,a_i),(a_i,a_1)}$ (see Lemma 6), we now infer that $vu^{(a_1,a_{i+1}),(a_{i+1},a_1)} \geq vu^{(a_1,a_i),(a_i,a_1)}$ for all $i \in \{1, \dots, m-1\}$. By chaining these inequalities, we derive that $vu^{(a_1,a_j),(a_j,a_1)} \geq vu^{(a_1,a_i),(a_i,a_1)}$ for all $i, j \in \{1, \dots, m\}$ with $i < j$. Finally, due to Lemma 7, this means that $vu^{(a_i,a_j),(a_j,a_i)} = vu^{(a_i,a_1),(a_1,a_i)} + vu^{(a_1,a_j),(a_j,a_1)} \geq 0$ for all $i, j \in \{1, \dots, m\}$ with $i < j$. Consequently, $vu^{(x,y),(y,x)} \geq 0$ for all $x, y \in X$ with $x \rhd y$ and thus $vu \geq 0$ for the vector $u = \sum_{(x,y)\in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$.

Furthermore, an analogous argument for $\rhd'$ shows that $vu^{(x,y),(y,x)} \geq 0$ for all $v \in \bar{D}_{\rhd'}$ and $x, y \in X$ with $x \rhd' y$. Hence, it also holds that $v \sum_{(x,y)\in \rhd' \setminus \rhd} u^{(x,y),(y,x)} \geq 0$ if $v \in \bar{D}_{\rhd'}$. Finally, we observe that $u = \sum_{(x,y)\in \rhd \setminus \rhd'} u^{(x,y),(y,x)} = -\sum_{(x,y)\in \rhd \setminus \rhd'} u^{(y,x),(x,y)} = -\sum_{(x,y)\in \rhd' \setminus \rhd} u^{(x,y),(y,x)}$ since $(x, y) \in \rhd \setminus \rhd'$ if and only if $(y, x) \in \rhd' \setminus \rhd$ and $u^{(x,y),(y,x)} = -u^{(y,x),(x,y)}$ for all $x, y \in A$. Thus, $vu = -v \sum_{(x,y)\in \rhd' \setminus \rhd} u^{(x,y),(y,x)} \leq 0$ if $v \in \bar{D}_{\rhd'}$ and u indeed separates $\bar{D}_{\rhd}$ from $\bar{D}_{\rhd'}$. □

We next turn to the case that U_2^X is linearly independent. In this case, we will use some of the ideas presented by Young and Levenglick (1978) to derive the vectors $u^{\rhd,\rhd'}$. In particular, just like Kemeny's rule, we note that every bivariate scoring rule can be seen as a linear optimization problem over a polytope. For formalizing this idea, we associate each ranking $\rhd \in \mathcal{R}(X)$ on some feasible set $X = \{a_1, \dots, a_{|X|}\}$ with a matrix $M^{\rhd} \in \mathbb{R}^{|X| \times |X|}$, which is defined by $M^{\rhd}_{i,j} = 1$ if $a_i \rhd a_j$ and $M^{\rhd}_{i,j} = 0$ otherwise. Moreover, let $\mathcal{M}$ denote the convex hull over the matrices $M^{\rhd}$, i.e., $M \in \mathcal{M}$ if and only if there are non-negative scalars $\lambda_{\rhd}$ for all $\rhd \in \mathcal{R}(X)$ such that $\sum_{\rhd \in \mathcal{R}(X)} \lambda_{\rhd} = 1$ and $M = \sum_{\rhd \in \mathcal{R}(X)} \lambda_{\rhd} M^{\rhd}$. Given a profile R and a ranking $\rhd \in \mathcal{R}(X)$, it holds that $\hat{s}(R, \rhd) = \sum_{a_i,a_j \in X \colon a_i \rhd a_j} \hat{s}(R, (a_i, a_j)) = \sum_{a_i,a_j \in X} M^{\rhd}_{i,j} \hat{s}(R, (a_i, a_j))$ for every bivariate scoring function s. Hence, each bivariate scoring rule f chooses the rankings that correspond to the optimal extreme points of the linear optimization problem that maximizes $\sum_{a_i,a_j \in X} M_{i,j} \hat{s}(R, (a_i, a_j))$ subject to $M \in \mathcal{M}$.

The reason why we are interested in this representation is the following basic fact of linear optimization: if an extreme point $M^{\rhd}$ does not maximize $\sum_{a_i,a_j \in X} M_{i,j} \hat{s}(R, (a_i, a_j))$, then there is a neighboring extreme point $M^{\rhd'}$ in $\mathcal{M}$ that

$$\begin{tikzcd}
\{a,c\} \arrow[r, leftrightarrow] \arrow[d, leftrightarrow] & \{a,d\} \arrow[d, leftrightarrow] \\
\{b,c\} \arrow[r, leftrightarrow] & \{b,d\}
\end{tikzcd}
\qquad
\begin{tikzcd}
\{a,d\} \arrow[d, leftrightarrow] \arrow[dr, leftrightarrow] & \{b,c\} \\
\{b,d\} & \{c,d\}
\end{tikzcd}
\qquad
\begin{tikzcd}
\{a,c\} \arrow[r, leftrightarrow] & \{c,d\}
\end{tikzcd}$$

$G(\rhd^1, \rhd^2)$ $\qquad$ $G(\rhd^1, \rhd^3)$ $\qquad$ $G(\rhd^2, \rhd^3)$

Figure 3: Transition graphs $G(\rhd^1, \rhd^2)$ (left), $G(\rhd^1, \rhd^3)$ (middle), and $G(\rhd^2, \rhd^3)$ (right) for the rankings $\rhd^1 = a, b, c, d$, $\rhd^2 = c, d, a, b$, and $\rhd^3 = d, a, c, b$.

achieves a higher objective value. To make this more formal, we say two rankings $\rhd, \rhd' \in \mathcal{R}(X)$ are neighbors of each other if the points $M^{\rhd}$ and $M^{\rhd'}$ are contained on a facet of dimension 1 of $\mathcal{M}$, and we define $\mathit{Neighbor}(\rhd) = \{\rhd' \in \mathcal{R}(X) \setminus \{\rhd\} \colon M^{\rhd}$ and $M^{\rhd'}$ are neighboring extreme points in $\mathcal{M}\}$ as the set of neighbors of $\rhd$. Based on this notion, we can compute bivariate scoring rules f only by considering the neighbors of every ranking, i.e., $f(R, X) = \{\rhd \in \mathcal{R}(X) \colon \forall \rhd' \in \mathit{Neighbor}(\rhd) \colon \hat{s}(R, \rhd) \geq \hat{s}(R, \rhd')\}$.

It will hence turn out sufficient to investigate the vectors $u^{\rhd,\rhd'}$ only for neighboring rankings $\rhd \in \mathcal{R}(X)$, $\rhd' \in \mathit{Neighbor}(\rhd)$. This leads to the question of when two rankings are neighbors in the polytope $\mathcal{M}$. To this end, we note that the polytope $\mathcal{M}$ is known as the permutation polytope, for which the neighborhood relation has been characterized by Gilmore and Hoffmann (1964) and Young (1978). To explain this characterization, we define the transposition graph $G(\rhd, \rhd')$ for two rankings $\rhd, \rhd'$ on some feasible set $X \in \mathcal{F}(A)$ by the vertex set $V = \{\{a, b\} \colon a, b \in X \wedge a \rhd b \wedge b \rhd' a\}$ and the edge set $E = \{\{\{a, b\}, \{b, c\}\} \colon a \neq c \wedge \{a, c\} \notin V \wedge \{a, b\}, \{b, c\} \in V\}$. To make this definition more accessible, consider two rankings $\rhd, \rhd' \in \mathcal{R}(X)$. The vertex set V of $G(\rhd, \rhd')$ is the set of pairs of alternatives on which $\rhd$ and $\rhd'$ disagree. Moreover, if $\{a, b\}, \{b, c\} \in V$ and $\{\{a, b\}, \{b, c\}\} \in E$ for some alternatives $a, b, c \in X$, then $\rhd$ and $\rhd'$ disagree on the order of $\{a, b\}$ and $\{b, c\}$, but agree on the order of $\{a, c\}$. This is the case if either *(i)* $a \rhd c \rhd b$ and $b \rhd' a \rhd' c$, *(ii)* $c \rhd a \rhd b$ and $b \rhd' c \rhd' a$, *(iii)* $b \rhd a \rhd c$ and $a \rhd' c \rhd' b$, or *(iv)* $b \rhd a \rhd c$ and $a \rhd' c \rhd' b$.

The characterization of Gilmore and Hoffmann (1964) then states that two rankings $\rhd, \rhd'$ are neighbors with respect to $\mathcal{M}$ if and only if $G(\rhd, \rhd')$ is connected. For example, for $X = \{a, b, c\}$ and $\rhd = a, b, c$, it can the be checked that $\mathit{Neighbor}(\rhd) = \{(a, c, b), (c, a, b), (b, a, c), (b, c, a)\}$. For a more involved example, we suppose that $X = \{a, b, c, d\}$ and let $\rhd^1 = a, b, c, d$, $\rhd^2 = c, d, a, b$, and $\rhd^3 = d, a, c, b$. The transposition graphs $G(\rhd^1, \rhd^2)$, $G(\rhd^1, \rhd^3)$, and $G(\rhd^2, \rhd^3)$ are shown in Figure 3. Based on these graphs and the characterization of Gilmore and Hoffmann (1964), it follows that $\rhd^1$ and $\rhd^2$ as well as $\rhd^2$ and $\rhd^3$ are neighbors as $G(\rhd^1, \rhd^2)$ and $G(\rhd^2, \rhd^3)$ are connected. By contrast, $\rhd^1$ and $\rhd^3$ are no neighbors as $G(\rhd^1, \rhd^3)$ is not connected.

Based on our new notation and insights, we will now analyze the vectors $u^{\rhd,\rhd'}$ for feasible sets X such that $|X| \geq 3$ and U_2^X is linearly independent.

Lemma 10. *Consider a feasible set $X \in \mathcal{F}(A)$ with $|X| \geq 3$ and two rankings $\rhd, \rhd' \in \mathcal{R}(X)$ with $\rhd' \in Neighbor(\rhd)$ and suppose that U_2^X is linearly independent. The vector $u = \sum_{(x,y)\in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ is non-zero and separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$.*

Proof. Consider an arbitrary feasible set $X \in \mathcal{F}(A)$ such that $|X| \geq 3$ and assume that U_2^X is linearly independent. Moreover, let $\rhd, \rhd' \in \mathcal{R}(X)$ denote two rankings such that $\rhd' \in Neighbor(\rhd)$. If $|\rhd \setminus \rhd'| = 1$, this lemma follows from Lemma 8 and the subsequent discussion, so we suppose that $|\rhd \setminus \rhd'| \geq 2$. In this case, we proceed in multiple steps to prove the lemma and define the vector $u^{\rhd,\rhd'} \in \mathbb{R}^{m!}$ as an arbitrary non-zero vector that separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$; such a vector exists by Lemma 4. First, we will show that the vector $u^{\rhd,\rhd'}$ is linearly dependent on U_2^X, which means that there are values λ_u such that $u^{\rhd,\rhd'} = \sum_{u\in U_2^X} \lambda_u u$. In our second step, we then prove that $\lambda_u = 0$ for all $u = u^{(x,y),(y,x)} \in U_2^X$ with $(x,y) \in \rhd$ if and only if $(x,y) \in \rhd'$. Since $u^{(x,y),(y,x)} = -u^{(y,x),(x,y)}$, this means that $u^{\rhd,\rhd'} = \sum_{(x,y)\in\rhd\setminus\rhd'} \lambda_{(x,y),(y,x)} u^{(x,y),(y,x)}$ for some values $\lambda_{(x,y),(y,x)}$. We then will show the lemma by proving that $u^{\rhd,\rhd'} = \lambda \sum_{(x,y)\in\rhd\setminus\rhd'} u^{(x,y),(y,x)}$ for some $\lambda > 0$ because this implies that $u = \sum_{(x,y)\in\rhd\setminus\rhd'} u^{(x,y),(y,x)}$ is a non-zero vector that separates $\bar{D}_\rhd$ from $\bar{D}_{\rhd'}$. In more detail, we will show this claim in Step 3 for the special case that $\rhd$ differs from $\rhd'$ by moving an alternative two positions down (or up). In Step 4, we then generalize this insight to arbitrary movements of a single alternative, and in Step 5, we complete the proof of the lemma.

Step 1: As the first step, we will show that the vector $u^{\rhd,\rhd'}$ is linearly dependent on U_2^X. To prove this claim, we assume for contradiction that it is not true. Now, consider the matrix M that contains all vectors in U_2^X and the vector $u^{\rhd,\rhd'}$ as rows. Since all rows in M are linearly independent, this matrix has full rank, so its image has full dimension. This means that there is a point $v \in \mathbb{R}^{m!}$ such that $vu^{(x,y),(y,x)} = 1$ for all $x, y \in X$ with $x \rhd y$ and $vu^{\rhd,\rhd'} = -1$. By the definition of $u^{\rhd,\rhd'}$, it follows that $v \notin \bar{D}_\rhd$. On the other hand, for every ranking $\rhd'' \neq \rhd$, there is a pair of alternatives $x, y \in X$ such that $x \rhd y$, $y \rhd'' x$, and $\rhd'' \in C(\{x,y\}, X)$.[8] By the choice of v and the definition of $u^{(x,y),(y,x)}$, it follows that $v \notin \bar{D}_{(y,x)}$ because $vu^{(x,y),(y,x)} = 1 > 0$. Moreover, since these alternatives are consecutive in $\rhd''$, Lemma 8 shows that $v \notin \bar{D}_{\rhd''}$, too. However, this means that $v \notin \bar{D}_{\rhd''}$ for any $\rhd'' \in \mathcal{R}(X)$, which contradicts that $\bigcup_{\rhd''\in\mathcal{R}(X)} \bar{D}_{\rhd''} = \mathbb{R}^{m!}$ (this is one of the basic properties of the sets $\bar{D}_{\rhd''}$; see the discussion after Lemma 3). This shows that our initial assumption was wrong, so $u^{\rhd,\rhd'}$ is linearly dependent of U_2^X.

Step 2: As a consequence of Step 1, there are values λ_u for all $u \in U_2^X$, not all of which are 0, such that $u^{\rhd,\rhd'} = \sum_{u\in U_2^X} \lambda_u u$. As the next step, we will show that $\lambda_u = 0$ for every vector $u = u^{(x,y),(y,x)} \in U_2^X$ such that $(x,y) \in \rhd$ if and only if $(x,y) \in \rhd'$. We assume for contradiction that this is not the case, which means that there are alternatives $a, b \in X$ such that either $(a,b) \in \rhd \cap \rhd'$ or $(a,b) \notin \rhd \cup \rhd'$ but $\lambda_{u^{(a,b),(b,a)}} \neq 0$ for $u^{(a,b),(b,a)} \in U_2^X$.

[8] To see this, let $\rhd = a_1, \dots, a_{|X|}$ and let i denote the smallest index such that $r(\rhd'', a_i) \neq i$; such an index exists as $\rhd \neq \rhd''$. By the choice of i, $r(\rhd'', a_j) = j = r(\rhd, a_j)$ for all $j < i$ and hence $r(\rhd'', a_i) > i$. This implies that there is an alternative a_k with $k > i$ such that $r(\rhd'', a_k) = r(\rhd'', a_i) - 1$, i.e., a_k and a_i are consecutive in $\rhd''$ and $a_k \rhd'' a_i$. Hence, we can set $x = a_i$ and $y = a_k$.

We subsequently focus on the case that $(a,b) \in \rhd \cap \rhd'$; if $(a,b) \notin \rhd \cup \rhd'$, we can replace $u^{(a,b),(b,a)}$ with $u^{(b,a),(a,b)} = -u^{(a,b),(b,a)}$ and $\lambda_{u^{(a,b),(b,a)}}$ with $-\lambda_{u^{(a,b),(b,a)}}$ in the presentation of $u^{\rhd,\rhd'}$ to arrive at the case that $(b,a) \in \rhd \cap \rhd'$.

Now, consider a vector $v \in \mathbb{R}^{m!}$ such that $vu = 0$ for all $u \in U_2^X \setminus \{u^{(a,b),(b,a)}\}$ and $vu^{(a,b),(b,a)} = 1$; such a vector exists as the set U_2^X is linearly independent. In more detail, we consider for this the matrix M that contains all vectors $u \in U_2^X$ as rows. Since U_2^X is linearly independent, the matrix M has full rank, which means that its image has full dimension. Consequently, the vector v indeed exists.[9] Note that we will use the same argument in the subsequent steps without going into further details.

Since $u^{\rhd,\rhd'} = \sum_{u\in U_2^X} \lambda_u u$ and $\lambda_{u^{(a,b),(b,a)}} \neq 0$, it holds that $vu^{\rhd,\rhd'} = \lambda_{u^{(a,b),(b,a)}} vu^{(a,b),(b,a)} \neq 0$. As the first case, we suppose that $vu^{\rhd,\rhd'} < 0$ and let $v' \in \mathbb{R}^{m!}$ denote a vector such that $v'u^{(x,y),(y,x)} > 0$ for all $x, y \in X$ with $x \rhd y$. Such a vector exists again due to the linear independence of U_2^X. Furthermore, let $\epsilon > 0$ be a sufficiently small value such that $v'' = v + \epsilon v'$ satisfies $v''u^{\rhd,\rhd'} < 0$. Since $v''u^{\rhd,\rhd'} < 0$, we infer that $v'' \notin \bar{D}_{\rhd}$. Next, for every ranking $\rhd'' \in \mathcal{R}(X) \setminus \{\rhd\}$, there is a consecutive pair of alternatives x, y with $x \rhd'' y$ and $y \rhd x$. By the definition of v'', it holds that $v''u^{(y,x),(x,y)} = \epsilon v'u^{(y,x),(x,y)} > 0$ if $\{x,y\} \neq \{a,b\}$ and $v''u^{(a,b),(b,a)} = 1 + \epsilon v'u^{(a,b),(b,a)} > 0$ as $a \rhd b$, so $v'' \notin \bar{D}_{(x,y)}$ and Lemma 8 entails that $v'' \notin \bar{D}_{\rhd''}$. However, this means that $v'' \notin \bar{D}_{\rhd''}$ for any $\rhd'' \in \mathcal{R}(X)$, which contradicts that $\bigcup_{\rhd''\in\mathcal{R}(X)} \bar{D}_{\rhd''} = \mathbb{R}^{m!}$. This is the desired contradiction, so $vu^{\rhd,\rhd'} < 0$ is impossible.

For the case that $vu^{\rhd,\rhd'} > 0$, we can simply replace $\rhd$ with $\rhd'$ in our argument. Then, we will construct a vector v'' such that $v'' \notin \bar{D}_{\rhd'}$ since $v''u^{\rhd,\rhd'} > 0$ and $v'' \notin \bar{D}_{\rhd''}$ for any other ranking $\rhd'' \in \mathcal{R}(X)$ because there are alternatives $x, y \in X$ such that x and y are consecutive in $\rhd''$, $x \rhd'' y$, and $v''u^{(y,x),(x,y)} > 0$. Hence, it is also not possible that $vu^{\rhd,\rhd'} > 0$, so our initial assumption that $\lambda_{u^{(a,b),(b,a)}} \neq 0$ is wrong.

Step 3: As the third step, we will prove the lemma for rankings $\rhd, \rhd'$ such that $\rhd$ differs from $\rhd'$ by only moving an alternative for two positions. We thus suppose that $\rhd = \dots, a, b, c, \dots$, and $\rhd' = \dots, b, c, a, \dots$, i.e., $\rhd'$ is derived from $\rhd$ by moving a two positions down (the case of moving an alternative up is symmetric). By the previous two steps, we know that $u^{\rhd,\rhd'} = \lambda_{(a,b),(b,a)} u^{(a,b),(b,a)} + \lambda_{(a,c),(c,a)} u^{(a,c),(c,a)}$ for some $\lambda_{(a,b),(b,a)}, \lambda_{(a,c),(c,a)} \in \mathbb{R}$. We will next show that $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)} > 0$. This implies that the vector $u = u^{(a,b),(b,a)} + u^{(a,c),(c,a)}$ is non-zero and that is separates $\bar{D}_{\rhd}$ from $\bar{D}_{\rhd'}$ as $u^{\rhd,\rhd'}$ satisfies both conditions.

To prove that $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)} > 0$, we consider a vector $v \in \mathbb{R}^{m!}$ such that $vu^{(a,b),(b,a)} = 1$, $vu^{(b,c),(c,b)} = 1$, $vu^{(c,a),(a,c)} = 1$, and $vu^{(x,y),(y,x)} = 1$ for all other pairs of alternatives $x, y \in X$ with $\{x,y\} \not\subseteq \{a,b,c\}$ and $x \rhd y$. As discussed in Step 2, such a vector exists due to the linear independence of U_2^X. Next, let τ denote the permutation defined by $\tau(a) = b$, $\tau(b) = c$, $\tau(c) = a$, and $\tau(x) = x$ for all other alternatives $x \in A \setminus$

[9] It may be helpful to exemplify this argument with Kemeny's rule. For this rule, the expression $vu^{(x,y),(y,x)}$ effectively computes the majority margin between x and y with respect to v (see Appendix A.3). Hence, the claim that the required vector v exists is equivalent to the fact that there is a (fractional) profile that has the corresponding majority margins. Our argument only generalizes this fact by observing that the linear independence of U_2^X implies the existence of such vectors v.

$\{a,b,c\}$. We will show that $\tau(v)u^{(x,y),(y,x)} = vu^{(x,y),(y,x)}$ for all distinct $x,y \in X$. First, if $x,y \in X\setminus\{a,b,c\}$, this follows from Lemma 6 since $\tau(u^{(x,y),(y,x)}) = u^{(\tau(x),\tau(y)),(\tau(y),\tau(x))} = u^{(x,y),(y,x)}$ and therefore $vu^{(x,y),(y,x)} = \tau(v)\tau(u^{(x,y),(y,x)}) = \tau(v)u^{(x,y),(y,x)}$. Next, suppose that $x,y \in \{a,b,c\}$; we will subsequently assume that $x = a$ and $y = b$ as all cases are symmetric. In this case, we recall that $vu^{(c,a),(a,c)} = 1$ and hence also $\tau(v)\tau(u^{(c,a),(a,c)}) = 1$. Since $\tau(u^{(c,a),(a,c)}) = u^{(a,b),(b,a)}$ by Lemma 6, it follows that $\tau(v)u^{(a,b),(b,a)} = 1 = vu^{(a,b),(b,a)}$. Finally, consider the case that $|\{x,y\} \cap \{a,b,c\}| = 1$; we suppose without loss of generality that $x = a$ and $y \notin \{a,b,c\}$. In this case, we note that $a \rhd y$ if and only if $c \rhd y$ as a, b, and c are consecutive in $\rhd$. Hence, $vu^{(a,y),(y,a)} = vu^{(c,y),(y,c)}$ by the definition of v. Since $\tau(u^{(c,y),(y,c)}) = u^{(a,y),(y,a)}$, we can conclude that $vu^{(a,y),(y,a)} = vu^{(c,y),(y,c)} = \tau(v)\tau(u^{(c,y),(y,c)}) = \tau(v)u^{(a,y),(y,a)}$. This completes the proof of our auxiliary claim. Furthermore, we note that an analogous argument shows that $\tau(\tau(v))u^{(x,y),(y,x)} = vu^{(x,y),(y,x)}$ for all distinct $x,y \in X$.

We now define the vector $v^* = \frac{1}{3}(v + \tau(v) + \tau(\tau(v)))$. For this vector, it holds that $v^*u^{(x,y),(y,x)} = vu^{(x,y),(y,x)}$ for all $x,y \in X$ due to our previous insights and that $\tau(v^*) = \frac{1}{3}(\tau(v) + \tau(\tau(v)) + \tau(\tau(\tau(v)))) = \frac{1}{3}(\tau(v) + \tau(\tau(v)) + v) = v^*$ since $\tau \circ \tau \circ \tau$ is the identity permutation. Our goal is to show that $v^* \in \bar{D}_{\rhd} \cap \bar{D}_{\rhd'}$ because this implies that $v^*u^{\rhd,\rhd'} = 0$. Since $v^*u^{(a,b),(b,a)} = 1$, $v^*u^{(a,c),(c,a)} = -v^*u^{(c,a),(a,c)} = -1$, and $v^*u^{\rhd,\rhd'} = \lambda_{(a,b),(b,a)}v^*u^{(a,b),(b,a)} + \lambda_{(a,c),(c,a)}v^*u^{(a,c),(c,a)}$, this proves that $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)}$.

Thus, consider an arbitrary ranking $\rhd''$ such that $v^* \in \bar{D}_{\rhd''}$; such a ranking exists as $\bigcup_{\rhd'' \in \mathcal{R}(X)} \bar{D}_{\rhd''} = \mathbb{R}^{m!}$. We next suppose that there is a pair of alternatives $x,y \in X$ with $\{x,y\} \not\subseteq \{a,b,c\}$ such that $x \rhd'' y$ and $y \rhd x$. In this case, we can also find a pair of alternatives x',y' such that x' and y' are consecutive in $\rhd''$, $\{x',y'\} \not\subseteq \{a,b,c\}$, $x' \rhd'' y'$, and $y' \rhd x'$.[10] However, $v^*u^{(y',x'),(x',y')} = 1$ by the definition of v^*, so $v^* \notin \bar{D}_{(x',y')}$ and Lemma 8 shows that $v^* \notin \bar{D}_{\rhd''}$ either. Consequently, if $v^* \in \bar{D}_{\rhd''}$, then $x \rhd'' y$ if and only if $x \rhd y$ for all $x,y \in X$ with $\{x,y\} \not\subseteq \{a,b,c\}$. Put differently, this means that $\rhd''$ can only disagree with $\rhd$ on the order on a, b, and c. Moreover, using again Lemma 8 and the vectors $u^{(a,b),(b,a)}$, $u^{(b,c),(c,b)}$, and $u^{(c,a),(a,c)}$, we get that $v^* \notin \bar{D}_{\rhd''}$ for $\rhd'' \in \{(\dots,b,a,c,\dots),(\dots,a,c,b,\dots),(\dots,c,b,a,\dots)\}$ (the symbol $\dots$ indicates here that the given rankings differ from $\rhd$ only in the ordering of a,b,c). Thus, $v^* \in \bar{D}_{\rhd''}$ can only be true for $\rhd'' \in \{\rhd,\rhd',\hat{\rhd}\}$, where $\hat{\rhd} = (\dots,c,a,b,\dots)$. Finally, due to the symmetry of v^* and the sets $\bar{D}_{\bar{\rhd}}$, it holds that $v^* = \tau(v^*)$, $\tau(\bar{D}_{\rhd}) = \bar{D}_{\rhd'}$, $\tau(\bar{D}_{\rhd'}) = \bar{D}_{\hat{\rhd}}$, and $\tau(\bar{D}_{\hat{\rhd}}) = \bar{D}_{\rhd}$. This implies that $v^* \in \bar{D}_{\rhd}$ if and only if $v^* \in \bar{D}_{\rhd'}$ if and only if $v^* \in \bar{D}_{\hat{\rhd}}$. Hence, v^* is in all three sets, so $v^* \in \bar{D}_{\rhd} \cap \bar{D}_{\rhd'}$ and $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)}$.

Since $u^{\rhd,\rhd'}$ is by definition a non-zero vector, $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)} \neq 0$. We will next show that $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)} > 0$. To this end, consider a point $v \in \mathbb{R}^{m!}$ such that $vu^{(x,y),(y,x)} = 1$ for all $x,y \in X$ with $x \rhd y$; in particular $vu^{(a,b),(b,a)} = vu^{(a,c),(c,a)} = 1$. Because there is for every ranking $\rhd'' \in \mathcal{R}(X)\setminus\{\rhd\}$ a consecutive pair of alternatives x,y

[10] To see this, we distinguish two cases: if $a \rhd b \rhd c \rhd x$, we consider the alternative $\bar{x}$ with $x \rhd'' \bar{x}$ that is consecutive to x in $\rhd''$. If $\bar{x} \rhd x$, we can set $x' = x$ and $y' = \bar{x}$ and are done. If $x \rhd \bar{x}$, it holds that $z \rhd x \rhd \bar{x}$ for all $z \in \{a,b,c,y\}$. We can hence repeat the argument with $x = \bar{x}$. Since this process always ensures that $x \rhd'' y$, it will eventually find a consecutive pair of alternatives x', y' that satisfies all requirements. If $x \in \{a,b,c\}$ or $x \rhd a \rhd b \rhd c$, it follows that $y \rhd a \rhd b \rhd c$ as $y \rhd x$ and $\{x,y\} \not\subseteq \{a,b,c\}$. Then, we can find x' and y' by moving from y upwards in $\rhd''$.

with $x \rhd'' y$ and $y \rhd x$, it follows that $v \notin \bar{D}_{\rhd''}$ for $\rhd'' \in \mathcal{R}(X) \setminus \{\rhd\}$ as $vu^{(y,x),(x,y)} = 1$. This means that $v \in \bar{D}_{\rhd}$ as otherwise $v \notin \bar{D}_{\rhd''}$ for every $\rhd'' \in \mathcal{R}(X)$. Consequently, $0 \leq vu^{\rhd,\rhd'} = \lambda_{(a,b),(b,a)} vu^{(a,b),(b,a)} + \lambda_{(a,c),(c,a)} vu^{(a,c),(c,a)} = \lambda_{(a,b),(b,a)} + \lambda_{(a,c),(c,a)}$. Since $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)} \neq 0$, this finally implies that $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)} > 0$, which completes the proof of this step.

Step 4: For our fourth step, we will generalize the insights of the last step to arbitrary shifts. For this, we assume that $\rhd = \dots, a, b_1, \dots, b_\ell, \dots$ and $\rhd' = \dots, b_1, \dots, b_\ell, a, \dots$, i.e., we derive $\rhd'$ from $\rhd$ by pushing down a for several positions. We note again that the case of moving up a is symmetric. If $\ell = 2$, we know from Step 3 that the vector $u^{(a,b_1),(b_1,a)} + u^{(a,b_2),(b_2,a)}$ separates $\bar{D}_{\rhd}$ from $\bar{D}_{\rhd'}$, so we suppose that $\ell \geq 3$. By Steps 1 and 2, we already have that $u^{\rhd,\rhd'} = \sum_{i \in \{1,\dots,\ell\}} \lambda_{(a,b_i),(b_i,a)} u^{(a,b_i),(b_i,a)}$ for some values $\lambda_{(a,b_1),(b_1,a)}, \dots, \lambda_{(a,b_\ell),(b_\ell,a)} \in \mathbb{R}$. Analogous to the last step, our goal is to show that $\lambda_{(a,b_i),(b_i,a)} = \lambda_{(a,b_j),(b_j,a)} > 0$ for all $i, j \in \{1, \dots, \ell\}$. We thus assume for contradiction that there is $i \in \{1, \dots, \ell - 1\}$ such that $\lambda_{(a,b_i),(b_i,a)} \neq \lambda_{(a,b_{i+1}),(b_{i+1},a)}$ and consider the cases that $\lambda_{(a,b_i),(b_i,a)} < \lambda_{(a,b_{i+1}),(b_{i+1},a)}$ and $\lambda_{(a,b_i),(b_i,a)} > \lambda_{(a,b_{i+1}),(b_{i+1},a)}$ separately.

Case 1: We first assume that $\lambda_{(a,b_i),(b_i,a)} < \lambda_{(a,b_{i+1}),(b_{i+1},a)}$, which means that there is an integer $k \in \mathbb{N}$ such that $k\lambda_{(a,b_{i+1}),(b_{i+1},a)} > k\lambda_{(a,b_i),(b_i,a)} + \sum_{j \in \{1,\dots,\ell\} \setminus \{i+1\}} \lambda_{(a,b_j),(b_j,a)}$. Now, consider a vector $v \in \mathbb{R}^{m!}$ such that $vu^{(a,b_i),(b_i,a)} = k + 1$, $vu^{(a,b_{i+1}),(b_{i+1},a)} = -k$, $vu^{(b_j,b_{j'}),(b_{j'},b_j)} = k + 1$ for all $j, j' \in \{1, \dots, \ell\}$ with $j < j'$, and $vu^{(x,y),(y,x)} = 1$ for all other alternatives $x, y \in X$ with $x \rhd y$. Such a vector exists as U_2^X is linearly independent. Our goal is to show that $v \notin \bar{D}_{\rhd''}$ for all $\rhd'' \in \mathcal{R}(X)$ as this contradicts that $\bigcup_{\rhd'' \in \mathcal{R}(X)} \bar{D}_{\rhd''} = \mathbb{R}^{m!}$.

To this end, we first note that $vu^{\rhd,\rhd'} = v \sum_{j \in \{1,\dots,\ell\}} \lambda_{(a,b_j),(b_j,a)} u^{(a,b_j),(b_j,a)} = -k\lambda_{(a,b_{i+1}),(b_{i+1},a)} + k\lambda_{(a,b_i),(b_i,a)} + \sum_{j \in \{1,\dots,\ell\} \setminus \{i+1\}} \lambda_{(a,b_j),(b_j,a)} < 0$, so $v \notin \bar{D}_{\rhd}$. Now, consider an arbitrary ranking $\rhd'' \neq \rhd$ and note that there is a consecutive pair of alternatives x, y in $\rhd''$ such that $x \rhd'' y$ and $y \rhd x$; otherwise $\rhd = \rhd''$. If $x \neq b_{i+1}$ or $y \neq a$, then it holds that $vu^{(y,x),(x,y)} > 0$. So, $v \notin \bar{D}_{(x,y)}$ by the definition of $u^{(y,x),(x,y)}$ and $v \notin \bar{D}_{\rhd''}$ by Lemma 8. Hence, the only consecutive pair in $\rhd''$ that is ordered differently than in $\rhd$ is (a, b_{i+1}), i.e., we have $b_{i+1} \rhd'' a$ and these two alternatives are consecutive in $\rhd''$. Next, let $X^+ = \{x \in X \colon x \rhd'' a\}$ and $X^- = \{x \in X \colon a \rhd'' x\}$ denote the alternatives that are ranked above and below a in $\rhd''$, respectively, and note that all alternatives in X^+ and X^- must be arranged as in $\rhd$ (i.e., $\rhd''|_{X^+} = \rhd|_{X^+}$ and $\rhd''|_{X^-} = \rhd|_{X^-}$). Otherwise, there is a pair of alternatives $x, y \in X^+$ (resp. $x, y \in X^-$) such that x, y are consecutive in $\rhd''$, $x \rhd'' y$, and $y \rhd x$; however, this contradicts that $v \in \bar{D}_{\rhd''}$ as $vu^{(y,x),(x,y)} > 0$. Moreover, analogous reasoning shows that there is no alternative $x \in X^-$ with $x \rhd a$ (because then $\rhd'' = \dots, b_{i+1}, a, x', \dots$ for some x' with $x' \rhd a$) and that there is no alternative $x \in X^+$ with $b_{i+1} \rhd x$ (because then $\rhd'' = \dots, x', a, \dots$ for some x' with $a \rhd b_{i+1} \rhd x'$).

Next, we suppose that there is an alternative $b_j \in X^-$ with $j \leq i$. Since the alternatives in X^- are ordered according to $\rhd$ and there is no alternative $x \in X^-$ with $x \rhd a$, this means that an alternative $b_{j'}$ with $j' \leq j \leq i$ is directly below a in $\rhd''$, i.e., $\rhd'' = \dots, b_{i+1}, a, b_{j'}, \dots$. Now, consider the ranking $\hat{\rhd} = \dots, a, b_{j'}, b_{i+1}, \dots$ that is derived from $\rhd''$ by pushing down b_{i+1} two positions. Due to Step 3, we can assume that $u^{\hat{\rhd},\rhd''} = u^{(a,b_{i+1}),(b_{i+1},a)} + u^{(b_{j'},b_{i+1}),(b_{i+1},b_{j'})}$, so we can compute that $vu^{\hat{\rhd},\rhd''} = vu^{(a,b_{i+1}),(b_{i+1},a)} + vu^{(b_{j'},b_{i+1}),(b_{i+1},b_{j'})} = -k+k+1 > 0$ because of the definition of v. This implies that $v \notin \bar{D}_{\rhd''}$. By contrast, if there is no alternative $b_j \in X^-$ with $j \leq i$, then $b_i \in X^+$. Moreover, since the alternatives in X^+ are ordered just as in $\rhd$, we have that $\rhd'' = \dots, b_i, b_{i+1}, a, \dots$. In this case, we consider the ranking $\hat{\rhd} = \dots, a, b_i, b_{i+1}, \dots$ and compute that $vu^{\hat{\rhd},\rhd''} = vu^{(a,b_i),(b_i,a)} + vu^{(a,b_{i+1}),(b_{i+1},a)} = k+1-k > 0$. It thus follows again that $v \notin \bar{D}_{\rhd''}$. Hence, we have shown that $v \notin \bar{D}_{\rhd''}$ for any ranking $\rhd'' \in \mathcal{R}(X)$, which contradicts that $\bigcup_{\rhd'' \in \mathcal{R}(X)} \bar{D}_{\rhd''} = \mathbb{R}^{m!}$. This shows that the assumption that $\lambda_{(a,b_i),(b_i,a)} < \lambda_{(a,b_{i+1}),(b_{i+1},a)}$ is wrong.

Case 2: For the second case, we suppose that $\lambda_{(a,b_i),(b_i,a)} > \lambda_{(a,b_{i+1}),(b_{i+1},a)}$. In this case, we will infer a contradiction by exchanging the roles of $\rhd$ and $\rhd'$. In more detail, we note that the vector $u^{\rhd',\rhd} = -u^{\rhd,\rhd'} = \sum_{j\in\{1,\dots,\ell\}} \lambda_{(a,b_j),(b_j,a)} u^{(b_j,a),(a,b_j)}$ separates $\bar{D}_{\rhd'}$ from $\bar{D}_{\rhd}$. Now, since $\lambda_{(a,b_i),(b_i,a)} > \lambda_{(a,b_{i+1}),(b_{i+1},a)}$, there is an integer $k \in \mathbb{N}$ such that $k\lambda_{(a,b_i),(b_i,a)} > k\lambda_{(a,b_{i+1}),(b_{i+1},a)} + \sum_{j\in\{1,\dots,\ell\}\setminus\{i\}} \lambda_{(a,b_j),(b_j,a)}$. Based on this integer k, we define the vector v by $vu^{(b_i,a),(a,b_i)} = -k$, $vu^{(b_{i+1},a),(a,b_{i+1})} = k+1$, $vu^{b_j,b_{j'}} = k+1$ for all $j, j' \in \{1,\dots,\ell\}$ with $j < j'$ and $vu^{(x,y),(y,x)} = 1$ for all other pairs of alternatives $x, y \in X$ with $x \rhd' y$; such a vector exists as U_2^X is linearly independent. It holds that $v \notin \bar{D}_{\rhd'}$ since $vu^{\rhd',\rhd} = -k\lambda_{(a,b_i),(b_i,a)} + k\lambda_{(a,b_{i+1}),(b_{i+1},a)} + \sum_{j\in\{1,\dots,\ell\}\setminus\{i\}} \lambda_{(a,b_j),(b_j,a)} < 0$.

Now, let $\rhd'' \neq \rhd'$ denote a ranking such that $v \in \bar{D}_{\rhd''}$ and define again the sets $X^+ = \{x \in X \colon x \rhd'' a\}$ and $X^- = \{x \in X \colon x \rhd'' a\}$. Since $\rhd'' \neq \rhd'$, there are two consecutive alternatives x, y in $\rhd''$ such that $x \rhd'' y$ and $y \rhd' x$. If $x \neq a$ or $y \neq b_i$, then it holds that $vu^{(y,x),(x,y)} > 0$ by the definition of v, which implies that $v \notin \bar{D}_{\rhd''}$ due to Lemma 8. This shows that $a \rhd'' b_i$ and that these two alternatives are consecutive in $\rhd''$. Moreover, analogous to the last case, we can infer that $\rhd''|_{X^+} = \rhd'|_{X^+}$, $\rhd''|_{X^-} = \rhd'|_{X^-}$, $x \notin X^+$ if $a \rhd' x$, and $x \notin X^-$ if $x \rhd' b_i$. This leaves us with two possible cases: first, if there is an alternative $b_j \in X^+$ with $j > i$, then $\rhd'' = \dots, b_{j'}, a, b_i, \dots$ for an alternative $b_{j'}$ with $j' \geq j > i$. This is true because there is no alternative $x \in X^+$ with $a \rhd' x$ and $\rhd''|_{X^+} = \rhd'|_{X^+}$. In this case, we let $\hat{\rhd} = \dots, b_i, b_{j'}, a, \dots$ be the ranking derived from $\rhd''$ by moving b_i up two positions. Since we can assume that $u^{\hat{\rhd},\rhd''} = u^{(b_i,a),(a,b_i)} + u^{(b_i,b_{j'}),(b_{j'},b_i)}$ due to Step 3, we compute that $vu^{\hat{\rhd},\rhd''} = vu^{(b_i,a),(a,b_i)} + vu^{(b_i,b_{j'}),(b_{j'},b_i)} = -k+k+1 > 0$. Hence, we conclude that $v \notin \bar{D}_{\rhd''}$. Next, if there is no alternative $b_j \in X^+$ with $j > i$, then $b_{i+1} \in X^-$ and $\rhd'' = \dots, a, b_i, b_{i+1}, \dots$. In this case, we consider the ranking $\hat{\rhd} = \dots, b_i, b_{i+1}, a, \dots$ derived from $\rhd''$ by moving a down two positions. Since $u^{\hat{\rhd},\rhd''} = u^{(b_i,a),(a,b_i)} + u^{(b_{i+1},a),(a,b_{i+1})}$ due to Step 3, it follows that $vu^{\hat{\rhd},\rhd''} = -k+k+1 > 0$, so $v \notin \bar{D}_{\rhd''}$. However, this means that $v \notin \bar{D}_{\rhd''}$ for any $\rhd'' \in \mathcal{R}(X)$, which contradicts that $\bigcup_{\rhd'' \in \mathcal{R}(X)} \bar{D}_{\rhd''} = \mathbb{R}^{m!}$.

Since both cases result in a contradiction, we conclude that $\lambda_{(a,b_i),(b_i,a)} = \lambda_{(a,b_{i+1}),(b_{i+1},a)}$ for all $i \in \{1,\dots,\ell-1\}$ and therefore also that $\lambda_{(a,b_i),(b_i,a)} = \lambda_{(a,b_j),(b_j,a)}$ for all $i, j \in \{1,\dots,\ell\}$. Moreover, $\lambda_{(a,b_i),(b_i,a)} \neq 0$ for all $i \in \{1,\dots,\ell\}$ since $u^{\rhd,\rhd'}$ is a non-zero vector. Finally, by considering the vector v with $vu^{(x,y),(y,x)} = 1$ for all $x, y \in X$

with $x \rhd y$, we can infer just as in Step 3 that $\lambda_{(a,b_i),(b_i,a)} > 0$ for all $i \in \{1, \dots, \ell\}$. Hence, $u^{\rhd,\rhd'} = \lambda \sum_{(x,y)\in\rhd\setminus\rhd'} u^{(x,y),(y,x)}$ for some $\lambda > 0$, which proves this step.

Step 5: As the last point, we consider arbitrary rankings $\rhd, \rhd' \in \mathcal{R}(X)$ such that $\rhd' \in \mathit{Neighbor}(\rhd)$ and we will show that $u^{\rhd,\rhd'} = \lambda \sum_{(x,y)\in\rhd\setminus\rhd'} u^{(x,y),(y,x)}$ for some $\lambda > 0$, which implies the lemma. To this end, we recall that $u^{\rhd,\rhd'} = \sum_{(x,y)\in\rhd\setminus\rhd'} \lambda_{(x,y),(y,x)} u^{(x,y),(y,x)}$ due to Steps 1 and 2. Hence, this step follows again by showing that all $\lambda_{(x,y),(y,x)}$ are equal and positive. For this, we will use the characterization of Gilmore and Hoffmann (1964) and Young (1978), which entails that the transposition graph $G(\rhd, \rhd')$ (defined before this lemma) is connected as $\rhd$ and $\rhd'$ are neighbors. Now, consider two connected vertices $\{a, b\}$ and $\{a, c\}$ in $G(\rhd, \rhd')$. Since these sets are connected vertices of $G(\rhd, \rhd')$, we know that $a \rhd b$ if and only if $b \rhd' a$, $a \rhd c$ if and only if $c \rhd' a$, and $b \rhd c$ if and only if $b \rhd' c$. This entails that $(\rhd|_{\{a,b,c\}}, \rhd'|_{\{a,b,c\}}) \in \{((a, b, c), (b, c, a)), ((a, c, b), (c, b, a)), ((b, c, a), (a, b, c)), ((c, b, a), (a, c, b))\}$. Since all these cases are symmetric, we suppose that $a \rhd b \rhd c$ and $b \rhd' c \rhd' a$ (this is not the same case as in Step 3 as there can be alternatives between a, b, and c in $\rhd$ and $\rhd'$). We next aim to show that $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)}$. Since $G(\rhd, \rhd')$ is connected and the vertices of this graph correspond to $\rhd \setminus \rhd'$, we can repeatedly apply this argument to infer that $\lambda_{(x,y),(y,x)} = \lambda_{(x',y'),(y',x')}$ for all pairs $(x, y), (x', y') \in \rhd \setminus \rhd'$.

To prove this claim, we consider a vector v with $vu^{(a,b),(b,a)} = vu^{(b,c),(c,b)} = vu^{(c,a),(a,c)} = 1$ and $vu^{(x,y),(y,x)} = 0$ for all other alternatives $x, y \in X$ with $\{x, y\} \not\subseteq \{a, b, c\}$; such a vector exists as U_2^X is linearly independent. Moreover, we let τ denote the permutation with $\tau(a) = b$, $\tau(b) = c$, $\tau(c) = a$, and $\tau(x) = x$ for all $x \in A \setminus \{a, b, c\}$. We will next show that $\tau(v)u^{(x,y),(y,x)} = vu^{(x,y),(y,x)}$ for all distinct $x, y \in X$. To prove this claim for $x, y \in X$ with $\{x, y\} \not\subseteq \{a, b, c\}$, we choose x' and y' such that $\tau(x') = x$ and $\tau(y') = y$. In particular, $\{x', y'\} \not\subseteq \{a, b, c\}$ as $\{x, y\} \not\subseteq \{a, b, c\}$. It therefore holds that $vu^{(x,y),(y,x)} = 0 = vu^{(x',y'),(y',x')} = \tau(v)\tau(u^{(x',y'),(y',x')}) = \tau(v)u^{(x,y),(y,x)}$. On the other hand, if $\{x, y\} \subseteq \{a, b, c\}$, we set $x = a$ and $y = b$ as all cases are symmetric. In this case, we note that $vu^{(a,b),(b,a)} = 1 = vu^{(c,a),(a,c)} = \tau(v)\tau(u^{(c,a),(a,c)}) = \tau(v)u^{(a,b),(b,a)}$. Moreover, an analogous argument also shows that $\tau(\tau(v))u^{(x,y),(y,x)} = vu^{(x,y),(y,x)}$ for all $x, y \in X$. Next, we define the vector v^* by $v^* = \frac{1}{3}(v + \tau(v) + \tau(\tau(v)))$ and observe that $\tau(v^*) = v^*$, $v^*u^{(a,b),(b,a)} = v^*u^{(b,c),(c,b)} = v^*u^{(c,a),(a,c)} = 1$, and $v^*u^{(x,y),(y,x)} = 0$ for all $x, y \in X$ with $\{x, y\} \not\subseteq \{a, b, c\}$. Our goal is to show that $v^* \in \bar{D}_\rhd \cap \bar{D}_{\rhd'}$ as this implies that $v^*u^{\rhd,\rhd'} = v^* \sum_{(x,y)\in\rhd\setminus\rhd'} \lambda_{(x,y),(y,x)} u^{(x,y),(y,x)} = \lambda_{(a,b),(b,a)} v^* u^{(a,b),(b,a)} + \lambda_{(a,c),(c,a)} v^* u^{(a,c),(c,a)} = \lambda_{(a,b),(b,a)} - \lambda_{(a,c),(c,a)} = 0$ and therefore that $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)}$.

To this end, we observe that $v^* \notin \bar{D}_{\rhd''}$ for any $\rhd''$ with $b \rhd'' a \rhd'' c$. To prove this claim, let $\hat{\rhd}$ denote the ranking that is derived from $\rhd''$ by moving a over b. By Step 4, the vector $\sum_{(x,y)\in\hat{\rhd}\setminus\rhd''} u^{(x,y),(y,x)}$ separates $\bar{D}_{\hat{\rhd}}$ from $\bar{D}_{\rhd''}$. In turn, the definition of v^* implies that $v^* \sum_{(x,y)\in\hat{\rhd}\setminus\rhd''} u^{(x,y),(y,x)} = v^* u^{(a,b),(b,a)} = 1$, so $v^* \notin \bar{D}_{\rhd''}$. A similar argument also rules out that $v^* \in \bar{D}_{\rhd''}$ for any $\rhd'' \in \mathcal{R}(X)$ with $c \rhd'' b \rhd'' a$ or $a \rhd'' c \rhd'' b$. Now, let $\rhd''$ denote a ranking with $v^* \in \bar{R}_{\rhd''}$; such a ranking exists since $\bigcup_{\rhd''\in\mathcal{R}(X)} \bar{D}_{\rhd''} = \mathbb{R}^{m!}$. Moreover, we assume that $a \rhd'' b \rhd'' c$ as all other cases are symmetric. Finally, let x, y denote two consecutive alternatives in $\rhd''$ such that $x \rhd'' y$ and $\{x, y\} \not\subseteq \{a, b, c\}$. Our goal is to show that $v^* \in \bar{D}_{\hat{\rhd}}$ for the ranking $\hat{\rhd}$ derived from $\rhd''$ by swapping x and y. By repeatedly applying this argument, we derive that $v^* \in \bar{D}_{\bar{\rhd}}$ for every ranking $\bar{\rhd}$ with $a \mathrel{\bar{\rhd}} b \mathrel{\bar{\rhd}} c$. Since $v^* = \tau(v^*) = \tau(\tau(v^*))$ and $\tau(v^*) \in \bar{D}_{\tau(\bar{\rhd})}$ if $v^* \in \bar{D}_{\bar{\rhd}}$, this entails that $v^* \in \bar{D}_{\bar{\rhd}}$ for all $\bar{\rhd}$ with $b \mathrel{\bar{\rhd}} c \mathrel{\bar{\rhd}} a$, which proves that $v^* \in \bar{D}_{\rhd} \cap \bar{D}_{\rhd'}$.

To prove this claim, we assume for contradiction that $v^* \not\in \bar{D}_{\hat{\rhd}}$. This means that there is another ranking $\rhd^*$ such that $v^* \hat{u}^{\hat{\rhd}, \rhd^*} < 0$ due to Lemma 5. Next, let v_1 denote a vector such that $v_1 u^{(b,c),(c,b)} = -1$, $v_1 u^{(c,a),(a,c)} = -2$, and $v_1 u^{(x,y),(y,x)} = 0$ for all other pairs of alternatives $x, y \in X$; such vectors exist as U_2^X is linearly independent. We define v_1^* as $v_1^* = v^* + \epsilon v_1$, where $\epsilon > 0$ is so small that $v^* u^{\rhd_1, \rhd_2} < 0$ implies $v_1^* u^{\rhd_1, \rhd_2} < 0$ for all rankings $\rhd_1, \rhd_2 \in \mathcal{R}(X)$. This means that if $v^* \not\in \bar{D}_{\rhd_1}$ for some ranking $\rhd_1$, then $v_1^* \not\in \bar{D}_{\rhd_1}$, too. Moreover, we claim that $v_1^* \not\in \bar{D}_{\rhd_1}$ for any ranking $\rhd_1$ with $b \rhd_1 c \rhd_1 a$ or $c \rhd_1 a \rhd_1 b$. For proving this, we first assume that $b \rhd_1 c \rhd_1 a$ and consider a ranking $\rhd_2$ which is derived from $\rhd_1$ by moving a over b. By Step 4, we can assume that $u^{\rhd_2, \rhd_1} = \sum_{(x,y) \in \rhd^2 \setminus \rhd^1} u^{(x,y),(y,x)}$, and the definition of v_1^* then implies that $v_1^* u^{\rhd_2, \rhd_1} = v_1^* u^{(a,b),(b,a)} + v_1^* u^{(a,c),(c,a)} = 1 - (1 - 2\epsilon) = 2\epsilon > 0$. Hence, $v^* \not\in \bar{D}_{\rhd_1}$. Similarly, if $c \rhd_1 a \rhd_1 b$, we consider the ranking $\rhd_2$ derived from $\rhd_1$ by moving c below b and infer that $v_1^* u^{\rhd_2, \rhd_1} = v_1^* u^{(a,c),(c,a)} + v_1^* u^{(b,c),(c,b)} = -(1 - 2\epsilon) + (1 - \epsilon) = \epsilon > 0$, so $v^* \not\in \bar{D}_{\rhd_1}$. This shows that v_1^* can only be in $\bar{D}_{\rhd_1}$ if $a \rhd_1 b \rhd_1 c$.

Next, let v_2 denote a vector such that $v_2 u^{(a,b),(b,a)} = v_2 u^{(b,c),(c,b)} = v_2 u^{(c,a),(a,c)} = 0$, and $v_2 u^{(x,y),(y,x)} = 1$ for all alternatives $x, y \in X$ with $\{x, y\} \not\subseteq \{a, b, c\}$ and $x \mathrel{\hat{\rhd}} y$; again, such a vector exists because of the linear independence of U_2^X. Moreover, let v_2^* be the vector defined by $v_2^* = v_1^* + \epsilon' v_2$, where $\epsilon' > 0$ is so small that $v_1^* u^{\rhd_1, \rhd_2} < 0$ implies $v_2^* u^{\rhd_1, \rhd_2} < 0$ for all $\rhd_1, \rhd_2 \in \mathcal{R}(X)$. Now, consider an arbitrary ranking $\rhd_1$ such that $v_2^* \in \bar{D}_{\rhd_1}$. By the choice of ϵ' and our previous insights, we get that $a \rhd_1 b \rhd_1 c$ and $\rhd_1 \neq \hat{\rhd}$ because $v^* \not\in \bar{D}_{\hat{\rhd}}$. Furthermore, since $\rhd_1 \neq \hat{\rhd}$, there is at least one consecutive pair of alternatives $x, y \in X$ such that $x \rhd_1 y$ and $y \mathrel{\hat{\rhd}} x$. Because $a \rhd_1 b \rhd_1 c$ and $a \mathrel{\hat{\rhd}} b \mathrel{\hat{\rhd}} c$, we also infer that $\{x, y\} \not\subseteq \{a, b, c\}$. Finally, it follows that $v_2^* u^{(x,y),(y,x)} = \epsilon' v_2 u^{(x,y),(y,x)} < 0$, so Lemma 8 shows that $v_2^* \not\in \bar{D}_{\rhd_1}$. However, this means that $v_2^* \not\in \bar{D}_{\bar{\rhd}}$ for any $\bar{\rhd} \in \mathcal{R}(X)$, which is a contradiction. Hence, $v^* \in \bar{D}_{\rhd''}$ implies $v^* \in \bar{D}_{\hat{\rhd}}$.

In turn, we now infer that $v^* \in \bar{D}_{\rhd_1}$ for all $\rhd_1 \in \mathcal{R}(X)$ with $a \rhd_1 b \rhd_1 c$. By the symmetry of v^* and our sets $\bar{D}_{\rhd_1}$, this means that $v^* \in \bar{D}_{\rhd_1}$ for all $\rhd_1 \in \mathcal{R}(X)$ with $b \rhd_1 c \rhd_1 a$, too. In particular, we thus have that $v^* \in \bar{D}_{\rhd} \cap \bar{D}_{\rhd'}$, so $v^* u^{\rhd_1, \rhd_2} = \lambda_{(a,b),(b,a)} - \lambda_{(a,c),(c,a)} = 0$. Hence, we derive that $\lambda_{(a,b),(b,a)} = \lambda_{(a,c),(c,a)}$. By applying this argument to all edges of the graph $G(\rhd, \rhd')$, it follows that $\lambda_{(x,y),(y,x)} = \lambda_{(x',y'),(y',x')}$ for all pairs of alternatives $(x, y), (x', y') \in \rhd \setminus \rhd'$. Next, we note that the $\lambda_{(x,y),(y,x)}$ cannot be 0 as $u^{\rhd, \rhd'}$ is a non-zero vector. Finally, by considering again the vector v with $v u^{(x,y),(y,x)} = 1$ for all $x, y \in X$ with $x \rhd y$, we infer that all $\lambda_{(x,y),(y,x)}$ are positive. Hence, $u^{\rhd, \rhd'} = \lambda \sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ for some $\lambda > 0$, which proves that the vector $\sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ is non-zero and separates $\bar{D}_{\rhd}$ from $\bar{D}_{\rhd'}$. $\square$

We are finally ready to prove our main theorem.

Theorem 1. *An SPF is a bivariate scoring rule if and only if it satisfies anonymity, neutrality, continuity, faithfulness, reinforcement, and local agenda consistency.*

Proof. Since the direction from right to left was proven in the main body, we focus on the converse. Thus, let f denote an SPF that satisfies all given axioms. First, by Lemma 3, there is a function $\hat{g}$ defined on $\mathbb{Q}^{m!} \times \mathcal{F}(A)$ that inherits all desirable properties of f and satisfies $f(R, X) = \hat{g}(v(R), X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$. Next, we define the sets $D_\rhd = \{v \in \mathbb{Q}^{m!} \colon \rhd \in \hat{g}(v, X)\}$ for every feasible set $X \in \mathcal{F}(A)$ and every ranking $\rhd \in \mathcal{R}(X)$. Moreover, we let $\bar{D}_\rhd$ denote the closure of $D_\rhd$ with respect to $\mathbb{R}^{m!}$. It follows that $\hat{g}(v, X) = \{\rhd \in \mathcal{R}(X) \colon v \in D_\rhd\} \subseteq \{\rhd \in \mathcal{R}(X) \colon v \in \bar{D}_\rhd\}$ for all $v \in \mathbb{Q}^{m!}$ and $X \in \mathcal{F}(A)$. Now, in Lemmas 3 to 5, we show that we can describe the sets $\bar{D}_\rhd$ by non-zero vectors $u^{\rhd,\rhd'}$ that satisfy $vu^{\rhd,\rhd'} \geq 0$ if $v \in \bar{D}_\rhd$ and $vu^{\rhd,\rhd'} \leq 0$ if $v \in \bar{D}_{\rhd'}$: it holds that $\bar{D}_\rhd = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathcal{R}(X) \setminus \{\rhd\} \colon vu^{\rhd,\rhd'} \geq 0\}$. Moreover, we suppose that the vectors $u^{(x,y),(y,x)}$ are those given in Lemma 6 for all distinct $x, y \in A$, and that $u^{\rhd,\rhd'} = \sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ for all $\rhd \in \mathcal{R}(X)$ and $\rhd' \in \mathit{Neighbor}(\rhd)$ (see Lemmas 9 and 10 for this).

Now, if $|X| = 2$, it holds that $\bar{D}_{(x,y)} = \{v \in \mathbb{R}^{m!} \colon vu^{(x,y),(y,x)} \geq 0\}$ since there are only the rankings (x, y) and (y, x) for $X = \{x, y\}$. Claim (4) of Lemma 6 then shows that there is a bivariate scoring function s such that $u_k^{(x,y),(y,x)} = s(r(\succ^k, x), r(\succ^k, y)) - s(r(\succ^k, y), r(\succ^k, x))$ for all distinct $x, y \in A$ and $k \in \{1, \dots, m!\}$. Consequently, it holds for all $v \in \mathbb{R}^{m!}$ and $x, y \in A$ that $vu^{(x,y),(y,x)} \geq 0$ if and only if $\hat{s}(v, (x, y)) \geq \hat{s}(v, (y, x))$, where $\hat{s}(v, (x, y)) = \sum_{k=1}^{m!} v_k s(r(\succ^k, x), r(\succ^k, y))$. This implies that $\bar{D}_{(x,y)} = \{v \in \mathbb{R}^{m!} \colon vu^{(x,y),(y,x)} \geq 0\} = \{v \in \mathbb{R}^{m!} \colon \hat{s}(v, (x, y)) \geq \hat{s}(v, (y, x))\}$. Thus, $\hat{g}(v, \{x, y\}) \subseteq \{\rhd \in \mathcal{R}(\{x, y\}) \colon \hat{s}(v, (x, y)) \geq \hat{s}(v, (y, x))\}$ for all $v \in \mathbb{Q}^{m!}$ and distinct $x, y \in A$.

For agendas X with $|X| \geq 3$, we define $\bar{D}_\rhd^N$ as $\bar{D}_\rhd^N = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathit{Neighbor}(\rhd) \colon vu^{\rhd,\rhd'} \geq 0\}$ and note that $\bar{D}_\rhd \subseteq \bar{D}_\rhd^N$. Since we assume that $u^{\rhd,\rhd'} = \sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)}$ for all $\rhd \in \mathcal{R}(X)$, $\rhd' \in \mathit{Neighbor}(\rhd)$, it follows that $\bar{D}_\rhd^N = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathit{Neighbor}(\rhd) \colon v \sum_{(x,y) \in \rhd \setminus \rhd'} u^{(x,y),(y,x)} \geq 0\}$. Moreover, it holds for all $v \in \mathbb{R}^{m!}$ that $\sum_{(x,y) \in \rhd \setminus \rhd'} vu^{(x,y),(y,x)} = \sum_{(x,y) \in \rhd \setminus \rhd'} \hat{s}(v, (x, y)) - \hat{s}(v, (y, x)) = \sum_{(x,y) \in \rhd} \hat{s}(v, (x, y)) - \sum_{(x,y) \in \rhd'} \hat{s}(v, (x, y))$. This means that $\bar{D}_\rhd^N = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathit{Neighbor}(\rhd) \colon \hat{s}(v, \rhd) \geq \hat{s}(v, \rhd')\}$, where $\hat{s}(v, \rhd) = \sum_{(x,y) \in \rhd} \hat{s}(v, (x, y))$.

Next, we note that $\bar{D}_\rhd^N = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathit{Neighbor}(\rhd) \colon \hat{s}(v, \rhd) \geq \hat{s}(v, \rhd')\} = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathcal{R}(X) \setminus \{\rhd\} \colon \hat{s}(v, \rhd) \geq \hat{s}(v, \rhd')\}$ because we can view the latter set as the solutions to a linear optimization problem. In more detail, we can associate every ranking $\rhd$ on a feasible set $X = \{a_1, \dots, a_\ell\}$ with the matrix $M^\rhd \in \mathbb{R}^{|X| \times |X|}$ defined by $M_{i,j}^\rhd = 1$ if $a_i \rhd a_j$ and $M_{i,j}^\rhd = 0$ otherwise. Furthermore, we define the set $\mathcal{M}$ as the convex hull over all $M^\rhd$ for $\rhd \in \mathcal{R}(X)$ and note that $\{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathcal{R}(X) \colon \hat{s}(v, \rhd) \geq \hat{s}(v, \rhd')\} = \{v \in \mathbb{R}^{m!} \colon M^\rhd \in \operatorname{argmax}_{M \in \mathcal{M}} \sum_{i,j \in \{1, \dots, \ell\}} M_{i,j} \hat{s}(v, (a_i, a_j))\}$ for every ranking $\rhd \in \mathcal{R}(X)$. It is a well-known fact in linear optimization that each non-optimal extreme point (the $M^\rhd$ in our case) has a neighbor with a higher objective value. Since the neighborhood relation of the matrices in $\mathcal{M}$ is equivalent to the neighborhood relationship of our rankings, we infer that $\{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathit{Neighbor}(\rhd) \colon \hat{s}(v, \rhd) \geq \hat{s}(v, \rhd')\} = \{v \in \mathbb{R}^{m!} \colon \forall \rhd' \in \mathcal{R}(X) \setminus \{\rhd\} \colon \hat{s}(v, \rhd) \geq \hat{s}(v, \rhd')\}$. Putting everything together, we derive for every profile $R \in \mathcal{R}^*$ and feasible set $X \in \mathcal{F}(A)$ that

$$\begin{aligned} f(R,X) = \hat{g}(v(R),X) &\subseteq \{\rhd \in \mathcal{R}(X) \colon v(R) \in \bar{D}^N_{\rhd}\} \\ &= \{\rhd \in \mathcal{R}(X) \colon \forall \rhd' \in \mathit{Neighbor}(\rhd) \colon \hat{s}(v(R),\rhd) \geq \hat{s}(v(R),\rhd')\} \\ &= \{\rhd \in \mathcal{R}(X) \colon \forall \rhd' \in \mathcal{R}(X) \colon \hat{s}(v(R),\rhd) \geq \hat{s}(v(R),\rhd')\}. \end{aligned}$$

This means that f returns always a subset of the bivariate scoring rule f' induced by s.

Finally, suppose for contradiction that $f(R,X) \subsetneq f'(R,X)$ for some profile $R \in \mathcal{R}^*$ and feasible set $X \in \mathcal{F}(A)$, and let $\rhd \in f'(R,X) \setminus f(R,X)$. Since f' satisfies all axioms of Lemma 2, there is a profile R' such that $f'(R',X) = \{\rhd\}$ and by our subset inclusion, we also have that $f(R',X) = \{\rhd\}$. In turn, by the reinforcement of f', it follows for every $\lambda \in \mathbb{N}$ that $f(\lambda R + R', X) = f'(\lambda R + R', X) = \{\rhd\}$. This, however, contradicts the continuity of f, which requires that there is a $\lambda \in \mathbb{N}$ such that $f(\lambda R + R', X) \subseteq f(R,X)$. Hence, $f(R,X) = f'(R,X)$ for all profiles $R \in \mathcal{R}^*$ and feasible sets $X \in \mathcal{F}(A)$, and f is the bivariate scoring rule induced by s. □